

(12)特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関  
国際事務局(43) 国際公開日  
2004 年 3 月 18 日 (18.03.2004)

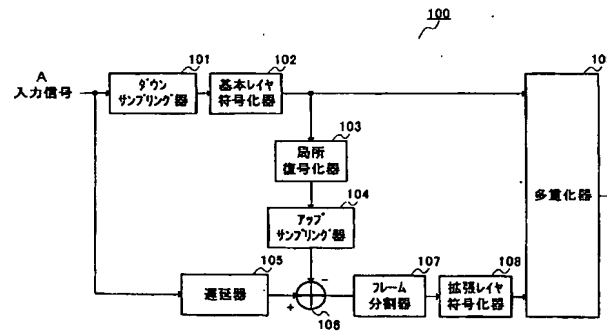
PCT

(10) 国際公開番号  
WO 2004/023457 A1

- (51) 国際特許分類: G10L 19/00 (72) 発明者; および  
(21) 国際出願番号: PCT/JP2003/010247 (75) 発明者/出願人 (米国についてのみ): 押切 正浩 (OSHIKIRI, Masahiro) [JP/JP]; 〒239-0841 神奈川県 横須賀市 野比1-31-7-A Kanagawa (JP).  
(22) 国際出願日: 2003 年 8 月 12 日 (12.08.2003)  
(25) 国際出願の言語: 日本語 (74) 代理人: 鷺田 公一 (WASHIDA, Kimihito); 〒206-0034 東京都 多摩市 鶴牧1丁目24-1 新都市センタービル 5 階 Tokyo (JP).  
(26) 国際公開の言語: 日本語  
(30) 優先権データ: 特願2002-261549 2002 年 9 月 6 日 (06.09.2002) JP (81) 指定国 (国内): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, (続業有)

(54) Title: SOUND ENCODING APPARATUS AND SOUND ENCODING METHOD

(54) 発明の名称: 音響符号化装置及び音響符号化方法



A...INPUT SIGNAL  
101...DOWN-SAMPLER  
102...BASIC LAYER ENCODER  
103...LOCAL DECODER  
104...UP-SAMPLER  
105...DELAYER  
106...SUBTRACTOR  
107...FRAME DIVIDER  
108...EXTENDED LAYER ENCODER  
109...MULTIPLEXER

(57) Abstract: A down-sampler (101) converts input data of a sampling rate FH to data of a sampling rate FL lower than the sampling rate FH. A basic layer encoder (102) encodes the input data of the sampling rate FL by predetermined unit of basic frame. A local decoder (103) decodes a first encoded code. An up-sampler (104) raises the sampling rate of the decoded signal to FH. A subtractor (106) subtracts the decoded signal from the input signal to provide a subtraction result as a residual signal. A frame divider (107) divides the residual signal into extended frames each shorter in time length than the basic frame. An extended layer encoder (108) encodes the residual signal that has been divided into the extended frames, and outputs, to a multiplexer (109), a second encoded code obtained by this encoding.

(57) 要約: ダウンサンプリング器 101 は、サンプリングレート FH の入力データをサンプリングレート FL より低いサンプリングレート FL に変換する。基本レイヤ符号化器 102 は、サンプリングレート FL の入力データを所定の基本フレーム単位で符号化する。局所復号化器 103 は、第 1 符号化コードを復号化する。アップサンプリング器 104 は、復号信号のサンプリングレートを FH に上げる。減算器 106 は、入力信号を復号信号で減算

(続業有)

BEST AVAILABLE COPY

WO 2004/023457 A1



SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC,  
VN, YU, ZA, ZM, ZW.

OAPI 特許 (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,  
ML, MR, NE, SN, TD, TG).

(84) 指定国 (広域): ARIPO 特許 (GH, GM, KE, LS, MW, MZ,  
SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア特許 (AM,  
AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ特許  
(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB,  
GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR),

添付公開書類:

— 国際調査報告書

2 文字コード及び他の略語については、定期発行される  
各 PCT ガゼットの巻頭に掲載されている「コードと略語  
のガイダンスノート」を参照。

## 明 細 書

## 音響符号化装置及び音響符号化方法

## 5 技術分野

本発明は、楽音信号または音声信号などの音響信号を高効率に圧縮符号化する音響符号化装置及び音響符号化方法に関し、特に符号化コードの一部からでも楽音や音声を復号することができるスケーラブル符号化を行う音響符号化装置及び音響符号化方法に関する。

10

## 背景技術

楽音信号または音声信号を低ビットレートで圧縮する音響符号化技術は、移動体通信における電波等の伝送路容量及び記録媒体の有効利用のために重要である。音声信号を符号化する音声符号化に、ITU (International  
15 Telecommunication Union)で規格化されているG726、G729などの方式がある。これらの方式は、狭帯域信号(300Hz～3.4kHz)を対象とし、8kbit/s～32kbit/sのビットレートで高品質に符号化できる。

また、広帯域信号(50Hz～7kHz)を符号化する標準方式としてITU  
20 UのG722、G722.1や、3GPP (The 3rd Generation Partnership Project)のAMR-WBなどが存在する。これらの方式は、6.6kbit/s～64kbit/sのビットレートで広帯域音声信号を高品質に符号化できる。

音声信号を低ビットレートで高効率に符号化を行う有効な方法に、CELP  
25 (Code Excited Linear Prediction)がある。CELPは、人間の音声生成モデルを工学的に模擬したモデルに基づき、乱数やパルス列で表される励振信号を周

期性の強さに対応するピッチフィルタと声道特性に対応する合成フィルタに通し、その出力信号と入力信号の二乗誤差が聴覚特性の重み付けの下で最小になるよう符号化パラメータを決定する方法である。（例えば、“Code-Excited Linear Prediction (CELP): high quality speech at very low bit rates”, Proc.

5 ICASSP 85, pp.937-940, 1985.参照)

最近の標準音声符号化方式の多くがCELPに基づいており、例えばG 7 2 9は、8 k b i t / s のビットレートで狭帯域信号の符号化でき、AMR-W Bは6. 6 k b i t / s ~ 2 3. 8 5 k b i t / s のビットレートで広帯域信号を符号化できる。

10 一方で、楽音信号を符号化する楽音符号化の場合、M P E G (Moving Picture Expert Group)で規格化されているレイヤ3方式やAAC方式のように、楽音信号を周波数領域に変換し、聴覚心理モデルを利用して符号化を行う変換符号化が一般的である。これらの方式は、サンプリングレートが4 4. 1 k H z の信号に対しチャネル当たり6 4 k b i t / s ~ 9 6 k b i t / s のビットレート  
15 トでほとんど劣化が生じないことが知られている。

しかしながら、音声信号が主体で、背景に音楽や環境音が重畳している信号を符号化する場合、音声符号化方式を適用すると背景部の音楽や環境音の影響で、背景部の信号のみならず音声信号も劣化してしまい全体的な品質が低下するという問題がある。これは、音声符号化方式が、CELPという音声モデル  
20 に特化した方式を基本にしているために生じる問題である。また、音声符号化方式が対応できる信号帯域は高々7 k H z までであり、それ以上の高域を持つ信号に対しては構成上十分に対応しきれないという問題がある。

一方で、楽音符号化は、音楽に対して高品質に符号化を行うことができるので、前述したような背景に音楽や環境音がある音声信号についても十分な品質  
25 を得ることができる。また、楽音符号化は、対象となる信号の帯域もCD品質であるサンプリングレートが2 2 k H z 程度の信号まで対応可能である。

その反面、高品質な符号化を実現するためにはビットレートを高くして使用する必要があり、仮にビットレートを  $32\text{ kbit/s}$  程度まで低く抑えると復号信号の品質が大きく低下するという問題がある。そのため、伝送レートの低い通信網で使用できないという問題がある。

- 5     上述した問題を回避するためにこれらの技術を組み合わせて、最初に入力信号を基本レイヤにて CELP で符号化し、次にその復号信号を入力信号から減算して得られる残差信号を求め、この信号を拡張レイヤにて変換符号化を行うスケーラブル符号化が考えられる。

- 10     この方法では、基本レイヤは CELP を用いているため音声信号を高品質に符号化でき、かつ拡張レイヤは基本レイヤで表しきれない背景の音楽や環境音、基本レイヤでカバーする周波数帯よりも高い周波数成分の信号を効率よく符号化することができる。さらにこの構成によればビットレートを低く抑えることができる。加えて、この構成によれば、符号化コードの一部つまり基本レイヤの符号化コードのみから音響信号を復号することが可能であり、このようなス  
15     ケーラブル機能は伝送容量の異なる複数のネットワークに対するマルチキャストの実現に有効である。

- しかしながら、このようなスケーラブル符号化では、拡張レイヤにて遅延が増大するという問題が生じる。この問題について図 1 及び図 2 を用いて説明する。図 1 は、従来の音声符号化における基本レイヤのフレーム(基本フレーム)  
20     と拡張レイヤのフレーム(拡張フレーム)の一例を示す図である。図 2 は、従来の音声復号化における基本レイヤのフレーム(基本フレーム)と拡張レイヤのフレーム(拡張フレーム)の一例を示す図である。

- 従来の音声符号化では、基本フレームと拡張フレームが、特定の同じ時間長のフレームで構成されている。図 1 において、時刻  $T(n-1) \sim T(n)$  に  
25     入力される入力信号は、第  $n$  基本フレームとなり基本レイヤにて符号化が行われる。これに対応して拡張レイヤでも時刻  $T(n-1) \sim T(n)$  の残差信号

に対して符号化が行われる。

ここで、拡張レイヤでMDCT(変形離散コサイン変換)を用いる場合、MDCTの分析フレームは前後に隣接する分析フレームと半分ずつ重ね合わせる必要がある。この重ね合わせは、合成時のフレーム間の不連続の発生を防ぐため

5    に行われる。

MDCTの場合、直交基底は分析フレーム内のみならず隣接する分析フレームとの間でも直交性が成り立つよう設計されており、そのために合成時に隣接する分析フレームと重ね合わせ加算することでフレーム間の不連続による歪の発生を防いでいる。図1では、第 $n$ 分析フレームは $T(n-2) \sim T(n)$ の  
10    長さに設定され、符号化処理が行われる。

復号化処理では、第 $n$ 基本フレームと第 $n$ 拡張フレームの復号信号が生成される。拡張レイヤではIMDCT(変形離散コサイン逆変換)が行われ、前述したように前フレーム(この場合は第 $n-1$ 拡張フレーム)の復号信号と合成フレーム長の半分だけ重ね合わせ加算を行う必要がある。そのために、復号化  
15    処理部では時刻 $T(n-1)$ の信号までしか生成することができない。

つまり、図2に示すような基本フレームと同じ長さの遅延(この場合は $T(n) - T(n-1)$ の時間長)が生じてしまう。仮に、基本フレームの時間長を20msとした場合、拡張レイヤで新たに生じる遅延は20msとなる。このような遅延の増大は、音声通話サービスを実現する上で深刻な問題となる。

20    このように、従来の装置においては、音声为主体で背景に音楽や雑音が重畳しているような信号を、遅延が短く低ビットレートで高品質に符号化を行うことが難しいという問題がある。

## 25    発明の開示

本発明の目的は、音声为主体で背景に音楽や雑音が重畳しているような信号

であっても、遅延が短く低ビットレートで高品質に符号化を行うことのできる音響符号化装置及び音響符号化方法を提供することである。

- この目的は、拡張レイヤのフレームの時間長を基本レイヤのフレームの時間長より短く設定して拡張レイヤの符号化を行い、音声为主体で背景に音楽や雑音
- 5 音が重畳しているような信号を遅延が短く低ビットレートで高品質に符号化を行うことにより達成される。

#### 図面の簡単な説明

- 図 1 は、従来の音声符号化における基本レイヤのフレーム(基本フレーム)と
- 10 拡張レイヤのフレーム(拡張フレーム)の一例を示す図、

図 2 は、従来の音声復号化における基本レイヤのフレーム(基本フレーム)と拡張レイヤのフレーム(拡張フレーム)の一例を示す図、

図 3 は、本発明の実施の形態 1 に係る音響符号化装置の構成を示すブロック図、

- 15 図 4 は、音響信号の情報の分布の一例を示す図、

図 5 は、基本レイヤと拡張レイヤで符号化の対象とする領域の一例を示す図、

図 6 は、基本レイヤと拡張レイヤの符号化の一例を示す図、

図 7 は、基本レイヤと拡張レイヤの復号化の一例を示す図、

- 図 8 は、本発明の実施の形態 1 に係る音響復号化装置の構成を示すブロック
- 20 図、

図 9 は、本発明の実施の形態 2 の基本レイヤ符号化器の内部構成の一例を示すブロック図、

図 10 は、本発明の実施の形態 2 の基本レイヤ復号化器の内部構成の一例を示すブロック図、

- 25 図 11 は、本発明の実施の形態 2 の基本レイヤ復号化器の内部構成の一例を示すブロック図、

図 1 2 は、本発明の実施の形態 3 の拡張レイヤ符号化器の内部構成の一例を示すブロック図、

図 1 3 は、M D C T 係数の配置の一例を示す図、

図 1 4 は、本発明の実施の形態 3 の拡張レイヤ復号化器の内部構成の一例を示すブロック図、

図 1 5 は、本発明の実施の形態 4 に係る音響符号化装置の構成を示すブロック図、

図 1 6 は、上記実施の形態の聴覚マスキング算出部の内部構成の一例を示すブロック図、

10 図 1 7 は、上記実施の形態の拡張レイヤ符号化器の内部構成の一例を示すブロック図、

図 1 8 は、上記実施の形態の聴覚マスキング算出部の内部構成の一例を示すブロック図、

15 図 1 9 は、本発明の実施の形態 5 の拡張レイヤ符号化器の内部構成の一例を示すブロック図、

図 2 0 は、M D C T 係数の配置の一例を示す図、

図 2 1 は、本発明の実施の形態 5 の拡張レイヤ復号化器の内部構成の一例を示すブロック図、

20 図 2 2 は、本発明の実施の形態 6 の拡張レイヤ符号化器の内部構成の一例を示すブロック図、

図 2 3 は、M D C T 係数の配置の一例を示す図、

図 2 4 は、本発明の実施の形態 6 の拡張レイヤ復号化器の内部構成の一例を示すブロック図、

図 2 5 は、本発明の実施の形態 7 に係る通信装置の構成を示すブロック図、

25 図 2 6 は、本発明の実施の形態 8 に係る通信装置の構成を示すブロック図、

図 2 7 は、本発明の実施の形態 9 に係る通信装置の構成を示すブロック図、



及び、

図 28 は、本発明の実施の形態 10 に係る通信装置の構成を示すブロック図である。

## 5 発明を実施するための最良の形態

以下、本発明の実施の形態について、図面を用いて説明する。

本発明者は、入力信号を符号化した基本フレームの時間長と、入力信号と符号化した入力信号を復号した信号との差分を符号化した拡張フレームの時間長が同一であることにより、復調時に長い遅延が発生することに着目し、本発明

## 10 をするに至った。

すなわち、本発明の骨子は、拡張レイヤのフレームの時間長を基本レイヤのフレームの時間長より短く設定して拡張レイヤの符号化を行い、音声为主体で背景に音楽や雑音を重ねているような信号を遅延が短く低ビットレートで高品質に符号化を行うことである。

## 15 (実施の形態 1)

図 3 は、本発明の実施の形態 1 に係る音響符号化装置の構成を示すブロック図である。図 3 の音響符号化装置 100 は、ダウンサンプリング器 101 と、基本レイヤ符号化器 102 と、局所復号化器 103 と、アップサンプリング器 104 と、遅延器 105 と、減算器 106 と、フレーム分割器 107 と、拡張  
20 レイヤ符号化器 108 と、多重化器 109 とから主に構成される。

図 3 において、ダウンサンプリング器 101 は、サンプリングレート FH の入力データ（音響データ）を受け付け、この入力データをサンプリングレート FH より低いサンプリングレート FL に変換して基本レイヤ符号化器 102 に出力する。

25 基本レイヤ符号化器 102 は、サンプリングレート FL の入力データを所定の基本フレーム単位で符号化し、入力データを符号化した第 1 符号化コードを

局所復号化器 103 と多重化器 109 に出力する。例えば、基本レイヤ符号化器 102 は、入力データを CELP 方式で符号化する。

局所復号化器 103 は、第 1 符号化コードを復号化し、復号化により得られた復号信号をアップサンプリング器 104 に出力する。アップサンプリング器

5 104 は、復号信号のサンプリングレートを FH に上げて減算器 106 に出力する。

遅延器 105 は、入力信号を所定の時間遅延して減算器 106 に出力する。

この遅延の大きさをダウンサンプリング器 101 と基本レイヤ符号化器 102 と局所復号器 103 とアップサンプリング器 104 で生じる時間遅れと同値と

10 することにより、次の減算処理での位相のずれを防ぐ役割を持つ。例えば、この遅延時間は、ダウンサンプリング器 101、基本レイヤ符号化器 102、局所復号化器 103、及びアップサンプリング器 104 における処理の時間の総和とする。減算器 106 は、入力信号を復号信号で減算し、減算結果を残差信号としてフレーム分割器 107 に出力する。

15 フレーム分割器 107 は、残差信号を基本フレームより時間長が短い拡張フレームに分割し、拡張フレームに分割した残差信号を拡張レイヤ符号化器 108 に出力する。拡張レイヤ符号化器 108 は、拡張フレームに分割された残差信号を符号化し、この符号化で得られた第 2 符号化コードを多重化器 109 に出力する。多重化器 109 は、第 1 符号化コードと第 2 符号化コードを多重化  
20 して出力する。

次に、本実施の形態に係る音響符号化装置の動作について説明する。ここでは、サンプリングレート FH の音響データである入力信号を符号化する例について説明する。

入力信号は、ダウンサンプリング器 101 において、サンプリングレート FH  
25 H より低いサンプリングレート FL に変換される。そして、サンプリングレート FL の入力信号は、基本レイヤ符号化器 102 において符号化される。そし

て、符号化された入力信号が局所復号化器 103 において復号化され、復号信号が生成される。復号信号は、アップサンプリング器 104 において、サンプリングレート FL より高いサンプリングレート FH に変換される。

一方、入力信号は、遅延器 105 において所定の時間遅延した後、減算器 106 5 106 に出力される。減算器 106 において遅延器 105 を介してきた入力信号とサンプリングレート FH に変換された復号信号との差分をとることにより、残差信号が得られる。

残差信号は、フレーム分割器 107 において、基本レイヤ符号化器 102 における符号化のフレーム単位より時間長の短いフレームに分割される。そして、  
10 分割された残差信号は、拡張レイヤ符号化器 108 において符号化される。基本レイヤ符号化器 102 において符号化された入力信号と、拡張レイヤ符号化器 108 において符号化された残差信号は、多重化器 109 において多重化される。

以下、基本レイヤ符号化器 102 と拡張レイヤ符号化器 108 とがそれぞれ  
15 符号化する信号について説明する。図 4 は、音響信号の情報の分布の一例を示す図である。図 4 において、縦軸は情報量を示し、横軸は周波数を示す。図 4 では、入力信号に含まれる音声情報と背景音楽・背景雑音情報がどの周波数帯にどれだけ存在しているかを表している。

図 4 に示すように、音声情報は、周波数の低い領域に情報が多く存在し、高  
20 域に向かうほど情報量は減少する。一方、背景音楽・背景雑音情報は、音声情報と比べると相対的に低域の情報は少なく、高域に含まれる情報が大きい。

そこで、基本レイヤでは CELP を用いて音声信号を高品質に符号化し、拡張レイヤでは基本レイヤで表しきれない背景の音楽や環境音、基本レイヤでカバーする周波数帯よりも高い周波数成分の信号を効率よく符号化する。

25 図 5 は、基本レイヤと拡張レイヤで符号化の対象とする領域の一例を示す図である。図 5 において、縦軸は情報量を示し、横軸は周波数を示す。図 5 は、

基本レイヤ符号化器 102 と拡張レイヤ符号化器 108 がそれぞれ符号化する情報の対象となる領域を表している。

基本レイヤ符号化器 102 は、0 ～ FL 間の周波数帯の音声情報を効率よく表すように設計されており、この領域での音声情報は品質良く符号化することができる。しかし、基本レイヤ符号化器 102 では、0 ～ FL 間の周波数帯の背景音楽・背景雑音情報の符号化品質が高くない。

拡張レイヤ符号化器 108 は、上記説明にある基本レイヤ符号化器 102 の能力不足の部分と、FL ～ FH 間の周波数帯の信号をカバーするように設計されている。よって、基本レイヤ符号化器 102 と拡張レイヤ符号化器 108 を組み合わせることで広い帯域で高品質な符号化が実現できる。

図 5 に示すように、基本レイヤ符号化器 102 における符号化により得られた第 1 符号化コードには、0 ～ FL 間の周波数帯の音声情報が含まれているので、少なくとも第 1 符号化コードのみでも復号信号が得られるというスケラブル機能が実現できる。

本実施の形態の音響符号化装置 100 では、この拡張レイヤ符号化器 108 において符号化するフレームの時間長を基本レイヤ符号化器 102 において符号化するフレームの時間長よりも十分に短く設定することにより、拡張レイヤで生じる遅延を短くする。

図 6 は、基本レイヤと拡張レイヤの符号化の一例を示す図である。図 6 において、横軸は時刻を示す。図 6 では、時刻  $T(n-1)$  から  $T(n)$  までの入力信号を第  $n$  フレームとして処理する。基本レイヤ符号化器 102 は、第  $n$  フレームを一つの基本フレームである第  $n$  基本フレームとして符号化を行う。一方、拡張レイヤ符号化器 108 は、第  $n$  フレームを複数の拡張フレームに分割して符号化する。

ここで、基本レイヤのフレーム（基本フレーム）に対して拡張レイヤのフレーム（拡張フレーム）の時間長は  $1/J$  に設定されている。図 6 では便宜上  $J$

= 8 に設定しているが、本実施例はこの数値に限定されることは無く、 $J \geq 2$  となる任意の整数を用いることができる。

図 6 の例では、 $J = 8$  としているので、拡張フレームが 8 個で基本フレーム 1 個に対応することになる。以後、第  $n$  基本フレームに対応する拡張フレーム  
5 のそれぞれを第  $n$  拡張フレーム (#  $j$ ) ( $j = 1 \sim 8$ ) と表記することにする。  
各拡張レイヤの分析フレームは、隣接するフレーム間で不連続が生じないように、分析フレームの半分が重なり合うように設定され、符号化処理が行われる。  
例えば、第  $n$  拡張フレーム (# 1) では、フレーム 401 とフレーム 402 をあ  
わせた領域が分析フレームとなる。そして、復号化側は、上記説明の入力信号  
10 を基本レイヤと拡張レイヤで符号化した信号を復号化する。

図 7 は、基本レイヤと拡張レイヤの復号化の一例を示す図である。図 7 に  
いて、横軸は時刻を示す。復号化処理では、第  $n$  基本フレームと第  $n$  拡張フ  
レームの復号信号が生成される。拡張レイヤでは、前フレームとの重ね合わせ加  
算が成立する区間の信号を復号することができる。図 7 では、時刻 501 まで、  
15 すなわち第  $n$  拡張フレーム (# 8) の中心の位置まで復号信号が生成される。

つまり、本実施の形態の音響符号化装置では、拡張レイヤで生じる遅延が時  
刻 501 から時刻 502 までであり、基本レイヤの時間長の  $1/8$  で済むこと  
になる。例えば、基本フレームの時間長が 20ms である場合、拡張レイヤで  
新たに生じる遅延は 2.5ms となる。

20 この例では、拡張フレームの時間長を基本フレームの時間長の  $1/8$  とした  
場合であったが、一般に拡張フレームの時間長を基本フレームの時間長の  $1/J$   
とした場合に、拡張レイヤで生じる遅延は  $1/J$  となり、本発明を適用する  
システムで許容される遅延の大きさによって  $J$  を設定することが可能である。

次に、上記復号化を行う音響復号化装置について説明する。図 8 は、本発明の  
25 実施の形態 1 に係る音響復号化装置の構成を示すブロック図である。図 8 の音  
響復号化装置 600 は、分離器 601 と、基本レイヤ復号化器 602 と、アッ

プサンプリング器603と、拡張レイヤ復号化器604と、重ね合わせ加算器605と、加算器606とから主に構成される。

分離器601は、音響符号化装置100において符号化されたコードを基本レイヤ用の第1符号化コードと拡張レイヤ用の第2符号化コードに分離し、第1符号化コードを基本レイヤ復号化器602に出力し、第2符号化コードを拡張レイヤ復号化器604に出力する。

基本レイヤ復号化器602は、第1符号化コードを復号してサンプリングレートFLの復号信号を得る。そして、基本レイヤ復号化器602は、復号信号をアップサンプリング器603に出力する。アップサンプリング器603は、サンプリングレートFLの復号信号をサンプリングレートFHの復号信号に変換して加算器606に出力する。

拡張レイヤ復号化器604は、第2符号化コードを復号してサンプリングレートFHの復号信号を得る。この第2符号化コードは、音響符号化装置100において、入力信号を基本フレームより時間長が短い拡張フレーム単位で符号化したコードである。そして、拡張レイヤ復号化器604は、この復号信号を重ね合わせ加算器605に出力する。

重ね合わせ加算器605は、拡張レイヤ復号化器604において復号された拡張フレーム単位の復号信号を重ね合わせ、重ね合わせた復号信号を加算器606に出力する。具体的には、重ね合わせ加算器605は、復号信号に合成用の窓関数を乗じ、前フレームで復号された時間領域の信号とフレームの半分だけオーバーラップさせて加算して出力信号を生成する。

加算器606は、アップサンプリング器603においてアップサンプリングされた基本レイヤの復号信号と、重ね合わせ加算器605において重ね合わされた拡張レイヤの復号信号とを加算して出力する。

このように、本実施の形態の音響符号化装置及び音響復号化装置によれば、音響符号化装置側で、基本フレームより短い時間長である拡張フレーム単位に

残差信号を分割し、分割した残差信号を符号化し、音響復号化装置側で、この基本フレームより短い時間長の拡張フレーム単位で符号化された残差信号を復号化し、時刻が重なる部分を重ね合わせることににより、復号化時の遅延の原因となる拡張フレームの時間長を短くすることができ、音声復号化の遅延を短く

5    することができる。

(実施の形態2)

本実施の形態では、基本レイヤの符号化においてCELPを用いる例について説明する。図9は、本発明の実施の形態2の基本レイヤ符号化器の内部構成の一例を示すブロック図である。図9は、図3の基本レイヤ符号化器102の  
10    内部構成を示す図である。図9の基本レイヤ符号化器102は、LPC分析器701と、聴感重み部702と、適応符号帳探索器703と、適応ベクトルゲイン量子化器704と、目標ベクトル生成器705と、雑音符号帳探索器706と、雑音ベクトルゲイン量子化器707と、多重化器708とから主に構成される。

15    LPC分析器701は、サンプリングレートFLの入力信号のLPC係数を算出し、このLPC係数をLSP係数などの量子化に適したパラメータに変換して量子化する。そして、LPC分析器701は、この量子化で得られる符号化コードを多重化器708に出力する。

また、LPC分析器701は、符号化コードから量子化後のLSP係数を算  
20    出してLPC係数に変換し、量子化後のLPC係数を、適応符号帳探索器703、適応ベクトルゲイン量子化器704、雑音符号帳探索器706、及び雑音ベクトルゲイン量子化器707に出力する。さらに、LPC分析器701は、量子化前のLPC係数を聴感重み部702に出力する。

聴感重み部702は、LPC分析器701で求められたLPC係数に基づい  
25    てダウンサンプリング器101から出力された入力信号に重み付けを行う。これは、量子化歪のスペクトルを入力信号のスペクトル包絡にマスクされるよう

スペクトル整形を行うことを目的としている。

適応符号帳探索器 703 では、聴覚重み付けされた入力信号を目標信号として適応符号帳の探索が行われる。過去の音源系列をピッチ周期で繰り返した信号を適応ベクトルと呼び、あらかじめ定められた範囲のピッチ周期で生成された適応ベクトルによって適応符号帳は構成される。

聴覚重み付けされた入力信号を  $t(n)$ 、ピッチ周期  $i$  の適応ベクトルに LPC 係数で構成される合成フィルタのインパルス応答を畳み込んだ信号を  $p_i(n)$  としたとき、適応符号帳探索器 703 は、式 (1) の評価関数  $D$  を最小とする適応ベクトルのピッチ周期  $i$  をパラメータとして多重化器 708 に出力する。

$$D = \sum_{n=0}^{N-1} t^2(n) - \frac{\left( \sum_{n=0}^{N-1} t(n)p_i(n) \right)^2}{\sum_{n=0}^{N-1} p_i^2(n)} \quad \dots (1)$$

ここで、 $N$  はベクトル長を表す。式 (1) の第 1 項はピッチ周期  $i$  に独立なので、実際には、適応符号帳探索器 703 は第 2 項のみを計算する。

適応ベクトルゲイン量子化器 704 は、適応ベクトルに乘じられる適応ベクトルゲインの量子化を行う。適応ベクトルゲイン  $\beta$  は、以下の式 (2) で表され、適応ベクトルゲイン量子化器 704 は、この適応ベクトルゲイン  $\beta$  をスカラー量子化し、量子化時に得られる符号を多重化器 708 に出力する。



$$\beta = \frac{\sum_{n=0}^{N-1} t(n) p_i(n)}{\sum_{n=0}^{N-1} p_i^2(n)} \quad \dots (2)$$

目標ベクトル生成器 705 は、入力信号から適応ベクトルの影響を減算して、雑音符号帳探索器 706 と雑音ベクトルゲイン量子化器 707 で用いる目標ベクトルを生成して出力する。目標ベクトル生成器 705 は、 $p_i(n)$  を式 1  
 5 で表される評価関数 D を最小とするときの適応ベクトルに合成フィルタのインパルス応答を畳み込んだ信号、 $\beta q$  を式 2 で表される適応ベクトル  $\beta$  をスカラー量子化したときの量子化値としたとき、目標ベクトル  $t_2(n)$  は、以下に示す式 (3) のように表される。

$$10 \quad t_2(n) = t(n) - \beta q \cdot p_i(n) \quad \dots (3)$$

雑音符号帳探索器 706 は、前記目標ベクトル  $t_2(n)$  と LPC 係数を用いて雑音符号帳の探索を行う。例えば、雑音符号帳探索器 706 には、ランダム雑音や大規模な音声信号を使って学習した信号を用いることができる。また、  
 15 雑音符号帳探索器 706 が備える雑音符号帳は、代数(Algebraic)符号帳のように、振幅 1 のパルスをあらかじめ定められた非常に少ない数だけ有するベクトルで表されることができる。この代数符号長は、パルスの位置とパルスの符号(極性)の最適な組み合わせを少ない計算量で決定することができるという特徴がある。

20 雑音符号帳探索器 706 は、目標ベクトルを  $t_2(n)$ 、コード  $j$  に対応する雑音ベクトルに合成フィルタのインパルス応答を畳み込んだ信号を  $c_j(n)$  としたとき、以下に示す式 (4) の評価関数 D を最小とする雑音ベクトルのイ

ンデックス  $j$  を多重化器 708 に出力する。

$$D = \sum_{n=0}^{N-1} t_2^2(n) - \frac{\left( \sum_{n=0}^{N-1} t_2(n) c_j(n) \right)^2}{\sum_{n=0}^{N-1} c_j^2(n)} \quad \dots (4)$$

- 5 雑音ベクトルゲイン量子化器 707 は、雑音ベクトルに乗じる雑音ベクトルゲインを量子化する。雑音ベクトルゲイン量子化器 707 は、以下に示す式(5)を用いて雑音ベクトルゲイン  $\gamma$  を算出し、この雑音ベクトルゲイン  $\gamma$  をスカラー量子化して多重化器 708 に出力する。

$$\gamma = \frac{\sum_{n=0}^{N-1} t_2(n) c_j(n)}{\sum_{n=0}^{N-1} c_j^2(n)} \quad \dots (5)$$

多重化器 708 は、送られてきた LPC 係数、適応ベクトル、適応ベクトルゲイン、雑音ベクトル、雑音ベクトルゲインの符号化コードを多重化して局所復号化器 103 及び多重化器 109 に出力する。

- 15 次に、復号化側について説明する。図 10 は、本発明の実施の形態 2 の基本レイヤ復号化器の内部構成の一例を示すブロック図である。図 10 は、図 8 の基本レイヤ復号化器 602 の内部構成を示す図である。図 10 の基本レイヤ復号化器 602 は、分離器 801 と、音源生成器 802 と、合成フィルタ 803 とから主に構成される。

- 20 分離器 801 は、分離器 601 から出力された第 1 符号化コードを LPC 係

数、適応ベクトル、適応ベクトルゲイン、雑音ベクトル、雑音ベクトルゲインの符号化コードに分離して、適応ベクトル、適応ベクトルゲイン、雑音ベクトル、雑音ベクトルゲインの符号化コードを音源生成器 802 に出力する。同様に、分離器 801 は、LPC 係数の符号化コードを合成フィルタ 803 に出力する。

音源生成器 802 は、適応ベクトル、適応ベクトルゲイン、雑音ベクトル、雑音ベクトルゲインの符号化コードを復号し、以下に示す式 (6) を用いて音源ベクトル  $ex(n)$  を生成する。

$$ex(n) = \beta_q \cdot q(n) + \gamma_q \cdot c(n) \quad \dots (6)$$

ここで、 $q(n)$  は適応ベクトル、 $\beta_q$  は適応ベクトルゲイン、 $c(n)$  は雑音ベクトル、 $\gamma_q$  は雑音ベクトルゲインを表す。

合成フィルタ 803 では、LPC 係数の符号化コードから LPC 係数を復号し、以下に示す式 (7) を用いて復号された LPC 係数から合成信号  $syn(n)$  を生成する。

$$syn(n) = ex(n) + \sum_{i=1}^{NP} \alpha_q(i) \cdot syn(n-i) \quad \dots (7)$$

ここで、 $\alpha_q$  は復号された LPC 係数、NP は LPC 係数の次数を表す。そして、合成フィルタ 803 は、復号された復号信号  $syn(n)$  をアップサンプリング器 603 に出力する。

このように、本実施の形態の音響符号化装置及び音響復号化装置によれば、

送信側において、基本レイヤにCELPを適用して入力信号を符号化し、受信側において、この符号化した入力信号にCELPを適用して復号することにより、低ビットレートで高品質な基本レイヤを実現することができる。

なお、本実施の形態の音声符号化装置は、量子化歪の知覚を抑制するために、  
5 合成フィルタ803の後にポストフィルタを従属接続する構成を採ることもできる。図11は、本発明の実施の形態2の基本レイヤ復号化器の内部構成の一例を示すブロック図である。但し、図10と同一の構成となるものについては、図10と同一番号を付し、詳しい説明を省略する。

ポストフィルタ901は、量子化歪の知覚の抑制の実現のために様々な構成  
10 を適用しうるが、代表的な方法として、分離器801で復号されて得られるLPC係数から構成されるホルマント強調フィルタを用いる方法がある。ホルマント強調フィルタ $H_f(z)$ は以下に示す式(8)で表される。

$$H_f(z) = \frac{A(z/\gamma_n)}{A(z/\gamma_d)} \cdot (1 - \mu z^{-1}) \quad \dots (8)$$

15

ここで、 $A(z)$ は復号LPC係数から構成される合成フィルタ、 $\gamma_n$ 、 $\gamma_d$ 、 $\mu$ はフィルタの特性を決定する定数を表す。

(実施の形態3)

本実施の形態の特徴は、拡張レイヤの入力信号を周波数領域の係数に変換し  
20 た後に符号化する変換符号化を用いる点にある。本実施の形態における拡張レイヤ符号化器108の基本構成を、図12を用いて説明する。図12は、本発明の実施の形態3の拡張レイヤ符号化器の内部構成の一例を示すブロック図である。図12は、図3の拡張レイヤ符号化器108の内部構成の一例を示す図である。図12の拡張レイヤ符号化器108は、MDCT部1001と、量子

化器 1002 とから主に構成される。

- MDCT 部 1001 は、フレーム分割器 107 から出力された入力信号を MDCT 変換(変形離散コサイン変換)して MDCT 係数を求める。MDCT 変換は、前後の隣接フレームと分析フレームを半分ずつ完全に重ね合わせ、分析フ
- 5 レームの前半部は奇関数、後半部は偶関数という直交基底を用いる。MDCT 変換は、波形を合成する際、逆変換後の波形を重ね合わせて加算することにより、フレーム境界歪が発生しないという特徴がある。MDCT を行う際には、sin 窓などの窓関数を入力信号に乗ずる。MDCT 係数を  $X(n)$  とすると、MDCT 係数は、以下に示す式 (9) に従い算出される。

10

$$X(m) = \sqrt{\frac{1}{N}} \sum_{n=0}^{2N-1} x(n) \cos \left\{ \frac{(2n+1+N) \cdot (2m+1)\pi}{4N} \right\} \quad \dots (9)$$

ここで  $X(n)$  は入力信号に窓関数を乗算した信号を表す。

- 量子化器 1002 は、MDCT 部 1001 で求められた MDCT 係数を量子
- 15 化する。具体的には、量子化器 1002 は、MDCT 係数それぞれをスカラー量子化する、または複数の MDCT 係数をまとめてベクトルとしベクトル量子化する。上記量子化方法は、特にスカラー量子化を適用する場合には、十分な品質を得るためにビットレートが高くなる傾向にある。そのため、この量子化方法は、拡張レイヤに十分なビットを配分することができる場合に有効である。
- 20 そして、量子化器 1002 は、MDCT 係数を量子化した符号を多重化器 109 に出力する。

次に、ビットレートの増加を抑えて効率よく MDCT 係数を量子化する方法について説明する。図 13 は、MDCT 係数の配置の一例を示す図である。図 13 において、横軸は時間、縦軸は周波数を表す。

拡張レイヤで符号化の対象となるMDCT係数は、図13で表されるように時間方向と、周波数方向の2次元のマトリクスで表すことができる。本実施の形態では1個の基本フレームに対し8個の拡張フレームを設定しているので横軸は8次元となり、縦軸は拡張フレームの長さに一致する次元数となる。図13では、縦軸を16次元で表しているが限定はなく、好ましくは時間を示す縦軸方向に60次元とするのが望ましい。

図13で表されるMDCT係数の全てについて十分高いSNRが得られるように量子化するには多くのビットが必要になる。この問題を回避するために、本実施の形態の音響符号化装置では、あらかじめ決めておいた帯域に含まれるMDCT係数のみを量子化し、それ以外のMDCT係数の情報は全く送らないようにする。つまり、図13の網掛け部分1101のMDCT係数を量子化し、それ以外のMDCT係数の量子化を行わないようにする。

この量子化方法は、基本レイヤが符号化の対象とする帯域(0~FL)は、既に基本レイヤで十分な品質で符号化されており十分な情報量を持つので、それ以外の帯域(例えばFL~FH)を拡張レイヤで符号化すれば良いという考えに基づく。もしくはこの量子化方法は、基本レイヤが符号化の対象としている帯域の高域部では符号化歪が大きくなりやすいという傾向があるため、基本レイヤが符号化の対象としている帯域の高域部と基本レイヤが符号化の対象としない帯域を符号化すれば良いという考えに基づく。

このように、基本レイヤの符号化でカバーできない領域のみ、もしくは基本レイヤの符号化でカバーできない領域と基本レイヤの符号化でカバーする帯域の一部を含む領域を符号化の対象とすることにより、符号化の対象となる信号を少なくすることができ、ビットレートの増加を抑えて効率よく変換係数を符号化することができる。

次に、復号化側について説明する。以下、周波数領域から時間領域への変換法に変形離散コサイン逆変換(IMDCT)を用いる場合について説明を行う。図

1 4は、本発明の実施の形態3の拡張レイヤ復号化器の内部構成の一例を示すブロック図である。図1 4は、図8の拡張レイヤ復号化器6 0 4の内部構成の一例を示す図である。図1 4の拡張レイヤ復号化器6 0 4は、MDC T係数復号化器1 2 0 1と、IMDC T部1 2 0 2とから主に構成される。

- 5 MDC T係数復号化器1 2 0 1は、分離器6 0 1から出力される第2符号化コードから量子化されたMDC T係数を復号する。IMDC T部1 2 0 2は、MDC T係数復号化器1 2 0 1から出力されるMDC T係数にIMDC Tを施し、時間領域の信号を生成して重ね合わせ加算器6 0 5に出力する。

- このように、本実施の形態の音響符号化装置及び音響復号化装置によれば、  
10 差分信号を時間領域から周波数領域に変換し、変換後の信号について基本レイヤの符号化によりカバーできない周波数領域を拡張レイヤで符号化することにより、音楽のようにスペクトルの変化が大きい信号にも対応することができる。

- なお、拡張レイヤが符号化の対象とする帯域をFL~FHに固定しなくても良い。基本レイヤの符号化方式の特性や入力信号の高域に含まれる情報量により拡張レイヤが効果的に機能する帯域が変わる。従って、実施の形態2で説明  
15 したように、基本レイヤに広帯域信号用のCELPを用い、さらに入力信号が音声である場合、拡張レイヤが符号化の対象とする帯域を6 kHz~9 kHzに設定すると良い。

(実施の形態4)

- 20 人間の聴覚特性には、ある信号が与えられたとき、その信号の周波数の近傍に位置する信号が聞こえなくなるというマスキング効果がある。本実施の形態の特徴は、入力信号を基に聴覚マスキングを求め、聴覚マスキングを利用して拡張レイヤの符号化を行う点にある。

- 図1 5は、本発明の実施の形態4に係る音響符号化装置の構成を示すブロック図である。ただし、図3と同一の構成となるものについては、図3と同一番号を付し、詳しい説明を省略する。図1 5の音響符号化装置1 3 0 0は、聴覚  
25

マスキング算出部 1301 と、拡張レイヤ符号化器 1302 とを具備し、マスキング効果の特性を利用して、入力信号のスペクトルから聴覚マスキングを算出し、量子化歪をこのマスキング値以下になるように MDCT 係数の量子化を行う点が図 3 の音響符号化装置と異なる。

- 5 遅延器 105 は、入力信号を所定の時間遅延して減算器 106 と聴覚マスキング算出部 1301 に出力する。聴覚マスキング算出部 1301 は、入力信号に基づいて、人間の聴覚では知覚できないスペクトルの大きさを示す聴覚マスキングを算出して拡張レイヤ符号化器 1302 に出力する。拡張レイヤ符号化器 1302 は、聴覚マスキングを超えるスペクトルを有する領域について差分  
10 信号を符号化して多重化器 109 に出力する。

- 次に、聴覚マスキング算出部 1301 の詳細について説明する。図 16 は、本実施の形態の聴覚マスキング算出部の内部構成の一例を示すブロック図である。図 16 の聴覚マスキング算出部 1301 は、FFT 部 1401 と、バークスペクトル算出器 1402 と、スプレッド関数畳み込み器 1403 と、トーン  
15 リティ算出器 1404 と、聴覚マスキング算出器 1405 とから主に構成される。

図 16 において、FFT 部 1401 は、遅延器 105 から出力された入力信号をフーリエ変換し、フーリエ係数  $\{ \text{Re}(m), \text{Im}(m) \}$  を算出する。ここで  $m$  は周波数を表す。

- 20 バークスペクトル算出器 1402 は、以下の式 (10) を用いてバークスペクトル  $B(k)$  を算出する。

$$B(k) = \sum_{m=f_l(k)}^{f_h(k)} P(m) \quad \dots (10)$$



ここで、 $P(m)$  はパワースペクトルを表し、以下の式 (11) より求められる。

$$P(m) = \text{Re}^2(m) + \text{Im}^2(m) \quad \dots (11)$$

5

ここで、 $\text{Re}(m)$ 、 $\text{Im}(m)$  はそれぞれ周波数  $m$  における複素スペクトルの実部および虚部を表す。また、 $k$  はバークスペクトルの番号に対応し、 $FL(k)$ 、 $FH(k)$  はそれぞれ第  $k$  バークスペクトルの最低周波数 (Hz)、最高周波数 (Hz) を表す。バークスペクトル  $B(k)$  はバークスケール上で等間隔に  
10 帯域分割されたときのスペクトル強度を表す。ヘルツスケールを  $f$ 、バークスケールを  $B$  と表したとき、ヘルツスケールとバークスケールの関係は以下の式 (12) で表される。

$$B = 13 \tan^{-1}(0.76f) + 3.5 \tan^{-1}\left(\frac{f}{7.5}\right) \quad \dots (12)$$

15 スプレッド関数畳み込み器 1403 は、バークスペクトル  $B(k)$  にスプレッド関数  $SF(k)$  を畳み込み、 $C(k)$  を算出する。

$$C(k) = B(k) * SF(k) \quad \dots (13)$$

20 トーナリティ算出器 1404 は、以下の式 (14) を用い、パワースペクトル  $P(m)$  から各バークスペクトルのスペクトル平坦度  $SFM(k)$  を求める。

$$SFM(k) = \frac{\mu g(k)}{\mu a(k)} \quad \dots (14)$$

- ここで、 $\mu g(k)$  は第  $k$  バークスペクトルの幾何平均、 $\mu a(k)$  は第  $k$  バークスペクトルの算術平均を表す。そして、トーンリティ算出器 1404 は、
- 5 以下の式 (15) を用いてスペクトル平坦度  $SFM(k)$  のデシベル値  $SFM_{dB}(k)$  からトーンリティ係数  $\alpha(k)$  を算出する。

$$\alpha(k) = \min\left(\frac{SFM_{dB}(k)}{-60}, 1.0\right) \quad \dots (15)$$

- 10 聴覚マスキング算出器 1405 は、以下の式 (16) を用いてトーンリティ算出器 1404 で算出したトーンリティ係数  $\alpha(k)$  から各バークスケールのオフセット  $O(k)$  を求める。

$$O(k) = \alpha(k) \cdot (14.5 - k) + (1.0 - \alpha(k)) \cdot 5.5 \quad \dots (16)$$

15

そして、聴覚マスキング算出器 1405 は、以下の式 (17) を用いてスプレッド関数畳み込み器 1403 で求めた  $C(k)$  からオフセット  $O(k)$  を減算して聴覚マスキング  $T(k)$  を算出する。

$$T(k) = \max\left(10^{\log_{10}(C(k)) - (O(k)/10)}, T_q(k)\right) \quad \dots (17)$$

20

ここで、 $T_q(k)$  は絶対閾値を表す。絶対閾値は、人間の聴覚特性として観測される聴覚マスキングの最小値を表す。そして、聴覚マスキング算出器 1 4 0 5 は、バークスケールで表される聴覚マスキング  $T(k)$  をヘルツスケール  $M(m)$  に変換して拡張レイヤ符号化器 1 3 0 2 に出力する。

このようにして求められた聴覚マスキング  $M(m)$  を使って、拡張レイヤ符号化器 1 3 0 2 にて MDCT 係数の符号化を行う。図 1 7 は、本実施の形態の拡張レイヤ符号化器の内部構成の一例を示すブロック図である。図 1 7 の拡張レイヤ符号化器 1 3 0 2 は、MDCT 部 1 5 0 1 と、MDCT 係数量子化器 1 5 0 2 とから主に構成される。

MDCT 部 1 5 0 1 は、フレーム分割器 1 0 7 から出力された入力信号に分析窓を乗じた後、MDCT 変換(変形離散コサイン変換)して MDCT 係数を求める。MDCT 変換は、前後の隣接フレームと分析フレームを半分ずつ完全に重ね合わせ、分析フレームの前半部は奇関数、後半部は偶関数という直交基底を用いる。MDCT 変換は、波形を合成する際、逆変換後の波形を重ね合わせて加算することにより、フレーム境界歪が発生しないという特徴がある。MDCT を行う際には、 $\sin$  窓などの窓関数を入力信号に乗ずる。MDCT 係数を  $X(n)$  とすると、MDCT 係数は、式 (9) に従い算出される。

MDCT 係数量子化器 1 5 0 2 は、MDCT 部 1 5 0 1 から出力された入力信号に聴覚マスキング算出部 1 3 0 1 から出力された聴覚マスキングを用いて入力信号を量子化する係数と量子化しない係数に分類し、量子化する係数のみを符号化する。具体的には、MDCT 係数量子化器 1 5 0 2 は、MDCT 係数  $X(m)$  と聴覚マスキング  $M(m)$  を比較し、 $M(m)$  よりも強度が小さい MDCT 係数  $X(m)$  はマスキング効果により人間の聴覚では知覚されないので無視して符号化の対象から外し、 $M(m)$  よりも強度の大きい MDCT 係数のみを量子化する。そして、MDCT 係数量子化器 1 5 0 2 は、量子化した MD

CT係数を多重化器109に出力する。

このように、本実施の形態の音響符号化装置によれば、マスキング効果の特性を利用して、入力信号のスペクトルから聴覚マスキングを算出し、拡張レイヤの符号化において、量子化歪をこのマスキング値以下になるように量子化を行うことにより、品質の劣化を伴わずに量子化の対象となるMDCT係数の数を減らすことができ、低ビットレートで高品質に符号化を行うことができる。

なお、上記実施の形態では、FFTを使った聴覚マスキングの算出法について説明しているが、FFTの代わりMDCTを使って聴覚マスキングを算出することもできる。図18は、本実施の形態の聴覚マスキング算出部の内部構成の一例を示すブロック図である。但し、図16と同一の構成となるものについては、図16と同一番号を付し、詳しい説明を省略する。

MDCT部1601は、MDCT係数を使ってパワースペクトル $P(m)$ を近似する。具体的には、MDCT部1601は、以下の式(18)を用いて $P(m)$ を近似する。

$$P(m) = R^2(m) \quad \dots (18)$$

ここで、 $R(m)$ は、入力信号をMDCT変換して求めたMDCT係数を表す。

バークスペクトル算出器1402は、MDCT部1601において近似された $P(m)$ からバークスペクトル $B(k)$ を算出する。それ以後は上述した方法に従い聴覚マスキングを算出する。

(実施の形態5)

本実施の形態は拡張レイヤ符号化器1302に関し、その特徴は聴覚マスキングを超えるMDCT係数を量子化の対象としたときに、MDCT係数の位置情報を効率よく符号化する方法に関するものである。

図19は、本発明の実施の形態5の拡張レイヤ符号化器の内部構成の一例を

示すブロック図である。図19は、図15の拡張レイヤ符号化器1302の内部構成の一例を示す図である。図19の拡張レイヤ符号化器1302は、MDCT部1701と、量子化位置決定部1702と、MDCT係数量子化器1703と、量子化位置符号化器1704と、多重化器1705とから主に構成される。

MDCT部1701は、フレーム分割器107から出力された入力信号に分析窓を乗じた後、MDCT変換(変形離散コサイン変換)してMDCT係数を求める。MDCT変換は、前後の隣接フレームと分析フレームを半分ずつ完全に重ね合わせ、分析フレームの前半部は奇関数、後半部は偶関数という直交基底を用いる。MDCT変換は、波形を合成する際、逆変換後の波形を重ね合わせて加算することにより、フレーム境界歪が発生しないという特徴がある。MDCTを行う際には、sin窓などの窓関数を入力信号に乗ずる。MDCT係数を $X(n)$ とすると、MDCT係数は、式(9)に従い算出される。

MDCT部1701で求められたMDCT係数を $X(j, m)$ と表す。ここで $j$ は拡張フレームのフレーム番号を表し、 $m$ は周波数を表す。本実施の形態では、拡張フレームの時間長を基本フレームの時間長の $1/8$ である場合について説明を行うものとする。図20は、MDCT係数の配置の一例を示す図である。MDCT係数 $X(j, m)$ は、図20に示すように横軸が時間、縦軸が周波数であるマトリクス上に表すことができる。MDCT部1701は、MDCT係数 $X(j, m)$ を量子化位置決定部1702とMDCT係数量子化器1703に出力する。

量子化位置決定部1702は、聴覚マスキング算出部1301から出力される聴覚マスキング $M(j, m)$ とMDCT部1701から出力されるMDCT係数 $X(j, m)$ を比較し、どの位置のMDCT係数を量子化の対象とすべきか決定する。

具体的には、量子化位置決定部1702は、以下の式(19)を満たす場合、

X (j, m) を量子化する。

$$|X(j, m)| - M(j, m) > 0 \quad \dots (19)$$

- 5      そして、量子化位置決定部 1702 は、以下の式 (20) を満たす場合、X (j, m) を量子化しない。

$$|X(j, m)| - M(j, m) \leq 0 \quad \dots (20)$$

- 10      そして、量子化位置決定部 1702 は、量子化の対象となる MDCT 係数 X (j, m) の位置情報を MDCT 係数量子化器 1703 と量子化位置符号化器 1704 に出力する。ここで、位置情報は、時間 j と周波数 m の組み合わせを指す。

- 図 20 では、量子化位置決定部 1702 で決定された量子化の対象となる MDCT 係数 X (j, m) の位置を網掛けで表している。この例では、(j, m) = (6, 1), (5, 3), ..., (7, 15), (5, 16) の位置にある MDCT 係数 X (j, m) が量子化の対象となる。

- 15      なお、ここで聴覚マスキング M (j, m) は拡張フレームに同期させて算出されているものとする。ただし計算量などの制限から、基本フレームに同期させて算出する構成でも良い。この場合、拡張フレームに同期させる場合に比べ聴覚マスキングの算出が 1/8 で済む。また、この場合、基本フレームで一度聴覚マスキングを求めた後に、同一の聴覚マスキングを全ての拡張フレームに対して使用することになる。

MDCT 係数量子化器 1703 は、量子化位置決定部 1702 で決定された

位置のMDCT係数 $X(j, m)$ を量子化する。量子化する際に、MDCT係数量子化器1703は、聴覚マスキング $M(j, m)$ の情報を利用し、量子化誤差が聴覚マスキング $M(j, m)$ 以下になるように量子化を行う。MDCT係数量子化器1703は、量子化後のMDCT係数を $X'(j, m)$ としたとき、以下の式(21)を満たすように量子化を行う。

$$|X(j, m) - X'(j, m)| \leq M(j, m) \quad \dots (21)$$

そして、MDCT係数量子化器1703は、量子化した後の符号を多重化器1705に出力する。

量子化位置符号化器1704は、位置情報を符号化する。例えば、量子化位置符号化器1704は、ランレングス法を適用して位置情報を符号化する。量子化位置符号化器1704は、周波数の低い方から時間軸方向に走査し、符号化の対象となる係数が連続して存在しない区間の数と符号の対象となる係数が連続して存在する区間の数を位置情報とする符号化を行う。

具体的には、 $(j, m) = (1, 1)$ から $j$ が増加する方向に走査し、符号化の対象となる係数があらわれるまでの座標の数を位置情報とする符号化を行う。そして、次に、符号化の対象となる係数までの座標の数をさらに位置情報とする。

図20では、 $(j, m) = (1, 1)$ から最初に符号化の対象となる係数の位置 $(j, m) = (1, 6)$ までの距離5、次に、符号化の対象となる係数は一つしか連続していないので1、次に符号化しない係数が連続する区間の数14となる。このように、図20では、位置情報を表す符号は、5、1、14、1、4、1、4・・・、5、1、3となる。量子化位置符号化器1704は、この位置情報を多重化器1705に出力する。多重化器1705は、MDCT

係数 $X(j, m)$ の量子化の情報と位置情報を多重化して多重化器109に出力する。

次に、復号化側について説明する。図21は、本発明の実施の形態5の拡張レイヤ復号化器の内部構成の一例を示すブロック図である。図21は、図8の  
5 拡張レイヤ復号化器604の内部構成の一例を示す図である。図21の拡張レイヤ復号化器604は、分離器1901と、MDCT係数復号化器1902と、量子化位置復号化器1903と、時間一周波数マトリクス生成器1904と、IMDCT部1905とから主に構成される。

分離器1901は、分離器601から出力された第2符号化コードをMDCT  
10 T係数量子化情報と量子化位置情報に分離し、MDCT係数量子化情報をMDCT係数復号化器1902に出力し、量子化位置情報を量子化位置復号化器1903に出力する。

MDCT係数復号化器1902は、分離器1901から出力されるMDCT  
15 係数量子化情報からMDCT係数を復号して時間一周波数マトリクス生成器1904に出力する。

量子化位置復号化器1903は、分離器1901から出力される量子化位置情報から量子化位置情報を復号して時間一周波数マトリクス生成器1904に出力する。この量子化位置情報は、復号MDCT係数のそれぞれが、時間一周波数マトリクスのどこに位置するかを表す情報である。

20 時間一周波数マトリクス生成器1904は、量子化位置復号化器1903から出力される量子化位置情報と、MDCT係数復号化器1902から出力される復号MDCT係数を用いて図20に示すような時間一周波数マトリクスを生成する。図20では、復号MDCT係数が存在する位置を網掛けで表し、復号MDCT係数が存在しない位置を白地で表している。白地の位置では復号MDCT  
25 CT係数が存在しないので、復号MDCT係数としてゼロが与えられる。

そして、時間一周波数マトリクス生成器1904は、各拡張フレーム( $j=1\sim$



J)毎に復号MDCT係数をIMDCT部1905に出力する。IMDCT部1905は、復号MDCT係数にIMDCTを施し、時間領域の信号を生成して重ね合わせ加算器605に出力する。

- このように、本実施の形態の音響符号化装置及び音響復号化装置によれば、
- 5 拡張レイヤにおける符号化において、残差信号を時間領域から周波数領域に変換した後、聴覚マスキングを行って符号化の対象となる係数を決定し、周波数とフレーム数の2次元での係数の位置情報を符号化することにより、符号化の対象となる係数と符号化の対象とならない係数の配置が連続することを利用して情報量を圧縮することができ、低ビットレートで高品質に符号化を行うこと
- 10 ができる。

(実施の形態6)

- 図2.2は、本発明の実施の形態6の拡張レイヤ符号化器の内部構成の一例を示すブロック図である。図2.2は、図1.5の拡張レイヤ符号化器1302の内部構成の一例を示す図である。但し、図1.9と同一の構成となるものについては、図1.9と同一番号を付し、詳しい説明を省略する。図2.2の拡張レイヤ符号化器1302は、領域分割器2001と、量子化領域決定部2002と、MDCT係数量子化器2003と、量子化領域符号化器2004とを具備し、聴覚マスキングを超えるMDCT係数を量子化の対象としたときに、MDCT係数の位置情報を効率よく符号化する別の方法に関するものである。
- 15 領域分割器2001は、MDCT部1701で求められたMDCT係数 $X(j, m)$ を複数の領域に分割される。ここでいう領域とは、複数のMDCT係数の位置をまとめたものを指し、符号化器と復号化器の両方に共通の情報としてあらかじめ定められたものである。

- 量子化領域決定部2002は、量子化の対象となる領域を決定する。具体的
- 25 には、量子化領域決定部2002は、領域を $S(k)$  ( $k=1 \sim K$ )と表したとき、領域 $S(k)$ に含まれるMDCT係数 $X(j, m)$ の内、このMDCT

係数  $X(j, m)$  が聴覚マスキング  $M(m)$  を超える量の総和を算出し、この総和の大きいものから  $K'$  個 ( $K' < K$ ) の領域を選択する。

図 23 は、MDCT 係数の配置の一例を示す図である。図 23 では、領域  $S(k)$  の一例を示している。図 23 の網掛け部は、量子化領域決定部 2002  
5 で決定された量子化の対象となる領域を表す。この例では、領域  $S(k)$  は時間軸方向に 4 次元、周波数軸方向に 2 次元の長方形になっており、量子化の対象は  $S(6)$ 、 $S(8)$ 、 $S(11)$ 、 $S(14)$  の 4 領域である。

量子化領域決定部 2002 は、前述したように MDCT 係数  $X(j, m)$  が聴覚マスキング  $M(j, m)$  を超える量の総和によってどの領域  $S(k)$  を量子化の対象とするか決定する。その総和  $V(k)$  は、以下の式 (22) より求  
10 められる。

$$V(k) = \sum_{(j,m) \in S(k)} (\text{MAX}(|X(j,m)| - M(j,m), 0))^2 \quad \dots (22)$$

15 この方法では、入力信号によっては高域の領域  $V(k)$  が選択されにくくなることもある。そこで、式 (22) の代わりに以下の式 (23) のような MDCT 係数  $X(j, m)$  の強度で正規化する方法を使用しても良い。

$$V(k) = \frac{\sum_{(j,m) \in S(k)} (\text{MAX}(|X(j,m)| - M(j,m), 0))^2}{\sum_{(j,m) \in S(k)} X(j,m)^2} \quad \dots (23)$$

20

そして、量子化領域決定部 2002 は、量子化の対象となる領域の情報を MDCT 係数量子化器 2003 と量子化領域符号化器 2004 に出力する。

量子化領域符号化器 2004 は、量子化の対象となる領域に符号 1、そうで

ない領域に符号0を割り振り、多重化器1705に出力する。図23の場合、符号は0000 0101 0010 0100となる。さらに、この符号をランレングスで表すことも可能である。その場合、得られる符号は5、1、1、1、2、1、2、1、2となる。

- 5 MDCT係数量子化器2003は、量子化領域決定部2002で決定された領域に含まれるMDCT係数の量子化を行う。量子化の方法としては、領域に含まれるMDCT係数から1つ以上のベクトルを構成し、ベクトル量子化を行う。ベクトル量子化の際、聴覚マスキング $M(j, m)$ で重み付けを行った尺度を用いても良い。
- 10 次に、復号化側について説明する。図24は、本発明の実施の形態6の拡張レイヤ復号化器の内部構成の一例を示すブロック図である。図24は、図8の拡張レイヤ復号化器604の内部構成の一例を示す図である。図24の拡張レイヤ復号化器604は、分離器2201と、MDCT係数復号化器2202と、量子化領域復号化器2203と、時間一周波数マトリクス生成器2204と、I
- 15 MDCT部2205とから主に構成される。

本実施の形態の特徴は、前述した実施の形態6の拡張レイヤ符号化器1302により生成された符号化コードを復号することができる点にある。

- 分離器2201は、分離器601から出力される第2符号化コードをMDCT係数量子化情報と量子化領域情報に分離し、MDCT係数量子化情報をMDCT係数復号化器2202に出力し、量子化領域情報を量子化領域復号化器2203に出力する。
- 20

MDCT係数復号化器2202は、分離器2201から得られるMDCT係数量子化情報からMDCT係数を復号する。量子化領域復号化器2203は、分離器2201から得られる量子化領域情報から量子化領域情報を復号する。

- 25 この量子化領域情報は、復号MDCT係数のそれぞれが、時間周波数マトリクスのどの領域に属するかを表す情報である。

時間一周波数マトリクス生成器 2204 は、量子化領域復号化器 2203 から得られる量子化領域情報と、MDCT 係数復号化器 2202 から得られる復号 MDCT 係数を使って図 23 に示すような時間一周波数マトリクスを生成する。図 23 では、復号 MDCT 係数が存在する領域を網掛けで表し、復号 MDCT 係数が存在しない領域を白地で表している。白地の領域では復号 MDCT 係数が存在しないので、復号 MDCT 係数としてゼロが与えられる。

そして、時間一周波数マトリクス生成器 2204 は、各拡張フレーム ( $j = 1 \sim J$ ) 毎に復号 MDCT 係数を IMDCT 部 2205 に出力する。IMDCT 部 2205 は、復号 MDCT 係数に IMDCT を施し、時間領域の信号を生成して重ね合わせ加算器 605 に出力する。

このように、本実施の形態の音響符号化装置及び音響復号化装置によれば、聴覚マスキングを超える残差信号が存在する時間領域と周波数領域の位置情報をグループ単位とすることにより、少ないビット数で符号化の対象となった領域の位置を表すことができるため、低ビットレート化を図ることができる。

#### 15 (実施の形態 7)

次に、本発明の実施の形態 7 について、図面を参照して説明する。図 25 は、本発明の実施の形態 7 に係る通信装置の構成を示すブロック図である。図 25 における信号処理装置 2303 は前述した実施の形態 1 から実施の形態 6 に示した音響符号化装置の中の 1 つによって構成されている点に本実施の形態の特徴がある。

図 25 に示すように、本発明の実施の形態 7 に係る通信装置 2300 は、入力装置 2301、A/D 変換装置 2302 及びネットワーク 2304 に接続されている信号処理装置 2303 を具備している。

A/D 変換装置 2302 は、入力装置 2301 の出力端子に接続されている。信号処理装置 2303 の入力端子は、A/D 変換装置 2302 の出力端子に接続されている。信号処理装置 2303 の出力端子はネットワーク 2304 に接

続されている。

入力装置 2301 は、人間の耳に聞こえる音波を電氣的信号であるアナログ信号に変換して A/D 変換装置 2302 に与える。A/D 変換装置 2302 はアナログ信号をディジタル信号に変換して信号処理装置 2303 に与える。信号処理装置 2303 は入力されてくるディジタル信号を符号化してコードを生成し、ネットワーク 2304 に出力する。

このように、本発明の実施の形態の通信装置によれば、通信において前述した実施の形態 1～6 に示したような効果を享受でき、少ないビット数で効率よく音響信号を符号化する音響符号化装置を提供することができる。

#### 10 (実施の形態 8)

次に、本発明の実施の形態 8 について、図面を参照して説明する。図 26 は、本発明の実施の形態 8 に係る通信装置の構成を示すブロック図である。図 26 における信号処理装置 2403 は前述した実施の形態 1 から実施の形態 6 に示した音響復号化装置の中の 1 つによって構成されている点に本実施の形態の特徴がある。

図 26 に示すように、本発明の実施の形態 8 に係る通信装置 2400 は、ネットワーク 2401 に接続されている受信装置 2402、信号処理装置 2403、及び D/A 変換装置 2404 及び出力装置 2405 を具備している。

受信装置 2402 の入力端子は、ネットワーク 2401 に接続されている。

20 信号処理装置 2403 の入力端子は、受信装置 2402 の出力端子に接続されている。D/A 変換装置 2404 の入力端子は、信号処理装置 2403 の出力端子に接続されている。出力装置 2405 の入力端子は、D/A 変換装置 2404 の出力端子に接続されている。

受信装置 2402 は、ネットワーク 2401 からのディジタルの符号化音響信号を受けてディジタルの受信音響信号を生成して信号処理装置 2403 に与える。信号処理装置 2403 は、受信装置 2402 からの受信音響信号を受け

てこの受信音響信号に復号化処理を行ってデジタルの復号化音響信号を生成してD/A変換装置2404に与える。D/A変換装置2404は、信号処理装置2403からのデジタルの復号化音声信号を変換してアナログの復号化音声信号を生成して出力装置2405に与える。出力装置2405は、電氣的信号であるアナログの復号化音響信号を空気の振動に変換して音波として人間の耳に聴こえるように出力する。

このように、本実施の形態の通信装置によれば、通信において前述した実施の形態1～6に示したような効果を享受でき、少ないビット数で効率よく符号化された音響信号を復号することができるので、良好な音響信号を出力することができる。

#### (実施の形態9)

次に、本発明の実施の形態9について、図面を参照して説明する。図27は、本発明の実施の形態9に係る通信装置の構成を示すブロック図である。本発明の実施の形態9において、図27における信号処理装置2503は、前述した実施の形態1から実施の形態6に示した音響符号化手段の中の1つによって構成されている点に本実施の形態の特徴がある。

図27に示すように、本発明の実施の形態9に係る通信装置2500は、入力装置2501、A/D変換装置2502、信号処理装置2503、RF変調装置2504及びアンテナ2505を具備している。

入力装置2501は人間の耳に聞こえる音波を電氣的信号であるアナログ信号に変換してA/D変換装置2502に与える。A/D変換装置2502はアナログ信号をデジタル信号に変換して信号処理装置2503に与える。信号処理装置2503は入力されてくるデジタル信号を符号化して符号化音響信号を生成し、RF変調装置2504に与える。RF変調装置2504は、符号化音響信号を変調して変調符号化音響信号を生成し、アンテナ2505に与える。アンテナ2505は、変調符号化音響信号を電波として送信する。

このように、本実施の形態の通信装置によれば、無線通信において前述した実施の形態 1 ～ 6 に示したような効果を享受でき、少ないビット数で効率よく音響信号を符号化することができる。

5      なお、本発明は、オーディオ信号を用いる送信装置、送信符号化装置又は音響信号符号化装置に適用することができる。また、本発明は、移動局装置又は基地局装置にも適用することができる。

(実施の形態 10)

次に、本発明の実施の形態 10 について、図面を参照して説明する。図 28 は、本発明の実施の形態 10 に係る通信装置の構成を示すブロック図である。  
10      本発明の実施の形態 10 において、図 28 における信号処理装置 2603 は、前述した実施の形態 1 から実施の形態 6 に示した音響復号化手段の中の 1 つによって構成されている点に本実施の形態の特徴がある。

図 28 に示すように、本発明の実施の形態 10 に係る通信装置 2600 は、アンテナ 2601、RF 復調装置 2602、信号処理装置 2603、D/A 変換装置 2604 及び出力装置 2605 を具備している。  
15      換装置 2604 及び出力装置 2605 を具備している。

アンテナ 2601 は、電波としてのデジタルの符号化音響信号を受けて電気信号のデジタルの受信符号化音響信号を生成して RF 復調装置 2602 に与える。RF 復調装置 2602 は、アンテナ 2601 からの受信符号化音響信号を復調して復調符号化音響信号を生成して信号処理装置 2603 に与える。

20      信号処理装置 2603 は、RF 復調装置 2602 からのデジタルの復調符号化音響信号を受けて復号化処理を行ってデジタルの復号化音響信号を生成して D/A 変換装置 2604 に与える。D/A 変換装置 2604 は、信号処理装置 2603 からのデジタルの復号化音声信号を変換してアナログの復号化音声信号を生成して出力装置 2605 に与える。出力装置 2605 は、電気的  
25      信号であるアナログの復号化音声信号を空気の振動に変換して音波として人間の耳に聴こえるように出力する。

このように、本実施の形態の通信装置によれば、無線通信において前述した実施の形態 1 ～ 6 に示したような効果を享受でき、少ないビット数で効率よく符号化された音響信号を復号することができるので、良好な音響信号を出力することができる。

- 5      なお、本発明は、オーディオ信号を用いる受信装置、受信復号化装置又は音声信号復号化装置に適用することができる。また、本発明は、移動局装置又は基地局装置にも適用することができる。

- また、本発明は上記実施の形態に限定されず、種々変更して実施することが可能である。例えば、上記実施の形態では、信号処理装置として行う場合につ  
10      いて説明しているが、これに限られるものではなく、この信号処理方法をソフトウェアとして行うことも可能である。

例えば、上記信号処理方法を実行するプログラムを予め ROM (Read Only Memory) に格納しておき、そのプログラムを CPU (Central Processor Unit) によって動作させるようにしても良い。

- 15      また、上記信号処理方法を実行するプログラムをコンピュータで読み取り可能な記憶媒体に格納し、記憶媒体に格納されたプログラムをコンピュータの RAM (Random Access memory) に記録して、コンピュータをそのプログラムにしたがって動作させるようにしても良い。

- 20      なお、上記説明では、時間領域から周波数領域への変換法に MDCT を用いる場合について説明を行っているがこれに限定されず直交変換であればいずれも適用できる。例えば、離散フーリエ変換または離散コサイン変換等を適用することもできる。

- 25      なお、本発明は、オーディオ信号を用いる受信装置、受信復号化装置又は音声信号復号化装置に適用することができる。また、本発明は、移動局装置又は基地局装置にも適用することができる。



以上の説明から明らかなように、本発明の音響符号化装置及び音響符号化方法によれば、拡張レイヤのフレームの時間長を基本レイヤのフレームの時間長より短く設定して拡張レイヤの符号化を行うことにより、音声为主体で背景に音楽や雑音が重畳しているような信号であっても、遅延が短く低ビットレート

5 で高品質に符号化を行うことができる。

本明細書は、2002年9月6日出願の特願2002-261549に基づくものである。この内容をここに含めておく。

#### 産業上の利用可能性

10 本発明は、楽音信号または音声信号などの音響信号を高能率に圧縮符号化する音響符号化装置及び通信装置に用いて好適である。

## 請求の範囲

1. 入力信号のサンプリングレートを下げるダウンサンプリング手段と、サンプリングレートが下げられた入力信号を所定の基本フレーム単位で符号化する基本レイヤ符号化手段と、符号化された入力信号を復号化して復号信号を得る復号化手段と、前記復号信号のサンプリングレートを入力時の入力信号のサンプリングレートと同一のレートに上げるアップサンプリング手段と、入力時の入力信号とサンプリングレートが上げられた復号信号との差分信号を得る減算手段と、前記基本フレームより時間長が短い拡張フレーム単位で前記差分信号を符号化する拡張レイヤ符号化手段と、を具備する音響符号化装置。
- 10 2. 基本フレーム単位の差分信号を拡張フレーム単位に分割するフレーム分割手段を具備し、前記拡張レイヤ符号化手段は分割された差分信号を符号化する請求の範囲第1項記載の音響符号化装置。
3. 前記基本レイヤ符号化手段は、符号励振線形予測法を用いて入力信号を符号化する請求の範囲第1項記載の音響符号化装置。
- 15 4. 前記拡張レイヤ符号化手段は、前記差分信号を時間領域から周波数領域に直交変換し、変換後の前記差分信号を符号化する請求の範囲第1項記載の音響符号化装置。
5. 前記拡張レイヤ符号化手段は、変形離散コサイン変換を用いて前記差分信号を時間領域から周波数領域に変換する請求の範囲第4項記載の音響符号化装置。
- 20 6. 前記拡張レイヤ符号化手段は、周波数領域に変換した前記差分信号を所定の帯域のみ符号化する請求の範囲第4項記載の音響符号化装置。
7. 聴覚に寄与しない振幅値を表す聴覚マスキングを算出する聴覚マスキング手段を具備し、前記拡張レイヤ符号化手段は、前記聴覚マスキング内の信号を符号化の対象としないことを特徴とする請求の範囲第4項記載の音響符号化装置。
- 25

8. 前記拡張レイヤ符号化手段は、聴覚マスキングと残差信号の差をとり、この差が相対的に大きい残差信号を符号化の対象とし、前記残差信号が存在する時間領域と周波数領域の位置を符号化する請求の範囲第7項記載の音響符号化装置。

- 5 9. 前記拡張レイヤ符号化手段は、時間領域と周波数領域の一方または両方において、それぞれ複数の領域を一つのグループとし、聴覚マスキングと残差信号の差を前記グループ単位で算出し、この差が相対的に大きいグループに含まれる残差信号のみを符号化する請求の範囲第8項記載の音響符号化装置。

- 10 10. 符号化側において入力信号を所定の基本フレーム単位で符号化した第1符号化コードを復号化して第1復号信号を得る基本レイヤ復号化手段と、入力信号と符号化側において第1符号化コードを復号した信号との残差信号を前記基本フレームより時間長が短い拡張フレーム単位で符号化した第2符号化コードを復号化して第2復号信号を得る拡張レイヤ復号化手段と、第1復号信号のサンプリングレートを前記第2復号信号のサンプリングレートと同一のサンプリングレートに上げるアップサンプリング手段と、前記第2復号信号とサンプリングレートを上げられた第1復号信号を加算する加算手段と、を具備する音響復号化装置。

11. 基本レイヤ復号化手段は、符号励振線形予測法を用いて第1符号化コードを復号化する請求の範囲第10項に記載の音響復号化装置。
- 20 12. 拡張レイヤ復号化手段は、第2符号化コードを復号化した信号を周波数領域から時間領域に直交変換する請求の範囲第10項に記載の音響復号化装置。
13. 第2復号信号同士を同じタイミングで符号化したフレーム部分を重ね合わせる重ね合わせ加算手段を具備し、前記拡張レイヤ復号化手段は、前記第2符号化コードを復号した信号に変形離散コサイン逆変換を用いて周波数領域から時間領域に直交変換することにより前記第2復号信号を復号して前記加算手段に出力し、加算手段は、前記重ね合わせ手段において重ね合わされた第2復

号信号と前記第1復号信号とを加算する請求の範囲第12項記載の音響復号化装置。

14. 前記拡張レイヤ復号化手段は、第2符号化コードから残差信号が存在する時間領域と周波数領域との情報を復号し、前記残差信号が存在する時間領域と周波数領域を復号化する請求の範囲第12項記載の音響復号化装置。

15. 前記拡張レイヤ復号化手段は、時間領域と周波数領域の一方または両方において、それぞれ複数の領域を一つのグループとし、復号の対象となるグループに含まれる残差信号を復号化する請求の範囲第14項記載の音響復号化装置。

- 10 16. 音響信号を電氣的信号に変換する音響入力手段と、この音響入力手段から出力された信号をデジタル信号に変換するA/D変換手段と、このA/D変換手段から出力されたデジタル信号を符号化する請求の範囲第1項記載の音響符号化装置と、この符号化装置から出力された符号化コードを無線周波数の信号に変調するRF変調手段と、このRF変調手段から出力された信号を電波に変換して送信する送信アンテナとを具備する音響信号送信装置。

17. 電波を受信する受信アンテナと、この受信アンテナに受信された信号を復調するRF復調手段と、このRF復調手段にて得られた情報を復号する請求の範囲第10項記載の音響復号化装置と、この復号化装置から出力された信号をアナログ信号に変換するD/A変換手段と、このD/A変換手段から出力された電氣的信号を音響信号に変換する音響出力手段とを具備する音響信号受信装置。

18. 請求の範囲第16項記載の音響信号送信装置記載の音響信号受信装置の少なくとも一方を具備する通信端末装置。
19. 請求の範囲第16項記載の音響信号送信装置記載の音響信号受信装置の少なくとも一方を具備する基地局装置。

20. 符号化側において、所定の基本フレーム単位で入力信号を符号化して第

- 1 符号化コードを作成し、符号化された前記入力信号を復号化して第1復号信号を得て、前記入力信号と前記復号信号との差分信号を得て、前記基本フレームより時間長が短い拡張フレーム単位で前記差分信号を符号化して第2符号化コードを作成し、復号化側において、前記第1符号化コードを復号化して第2
- 5 復号信号を得て、前記第2符号化コードを復号化して第3復号信号を得て、前記第2復号信号と前記第3復号信号を加算する音響符号化方法。

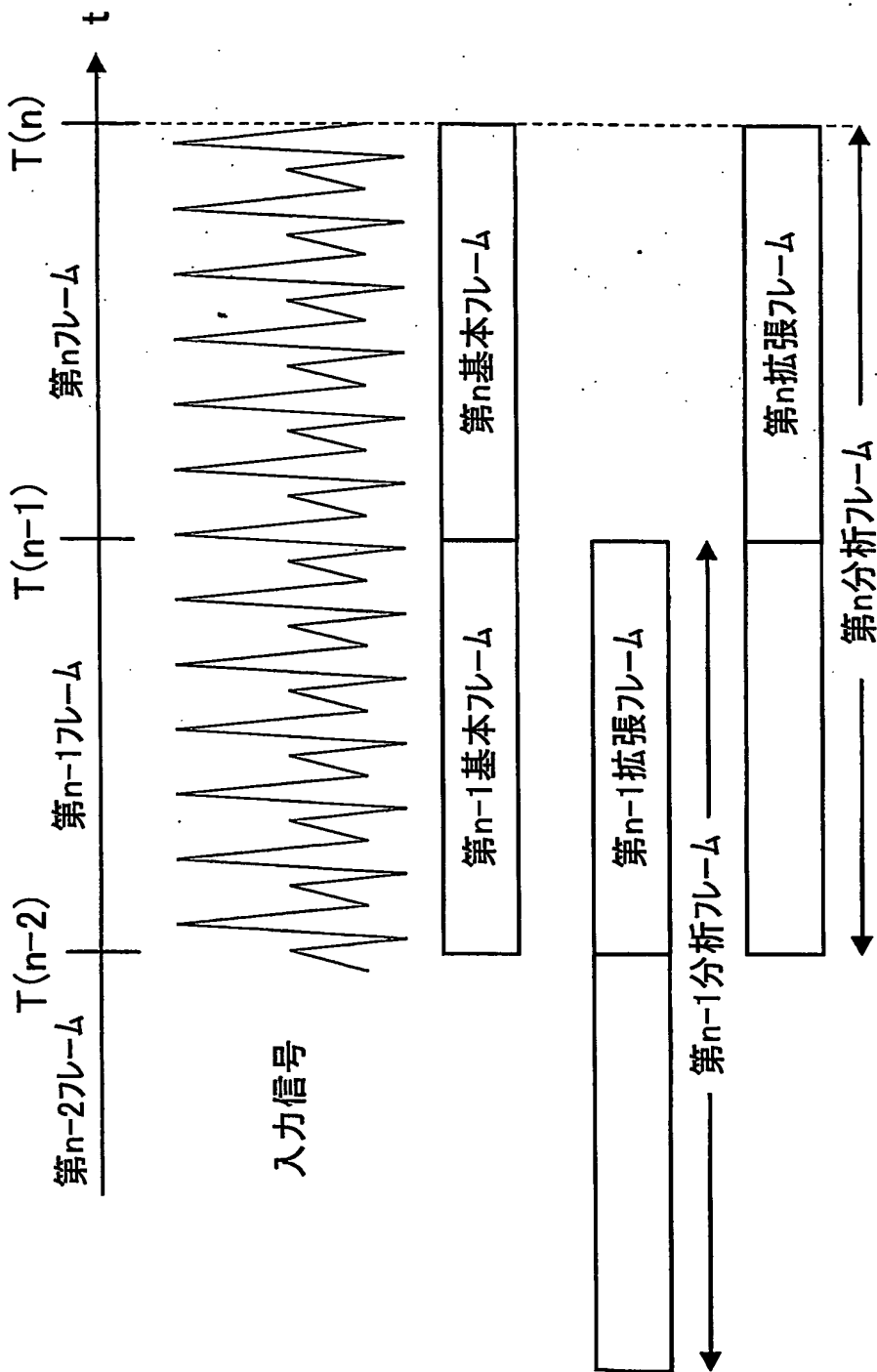


図1

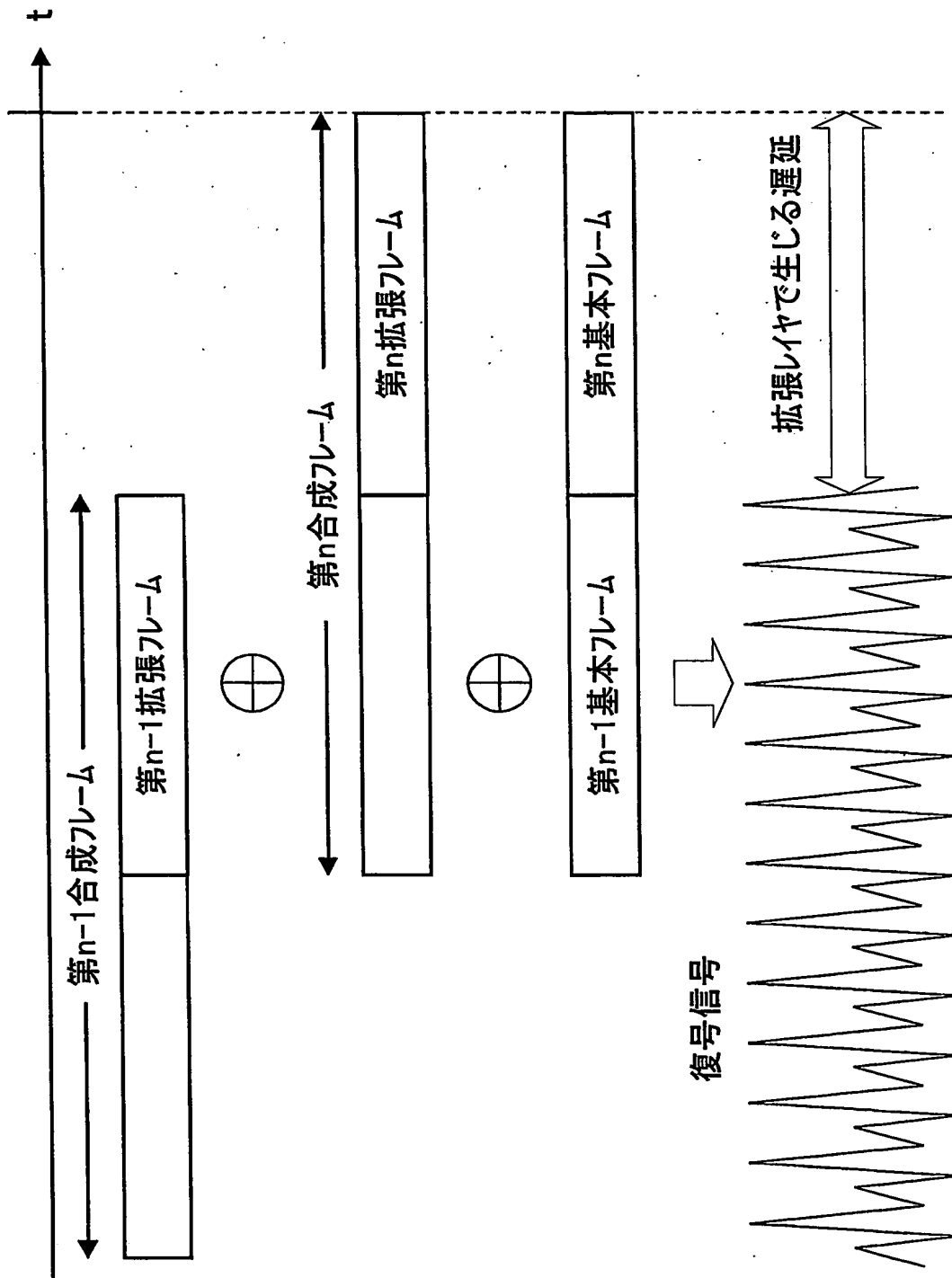


図2

3/26

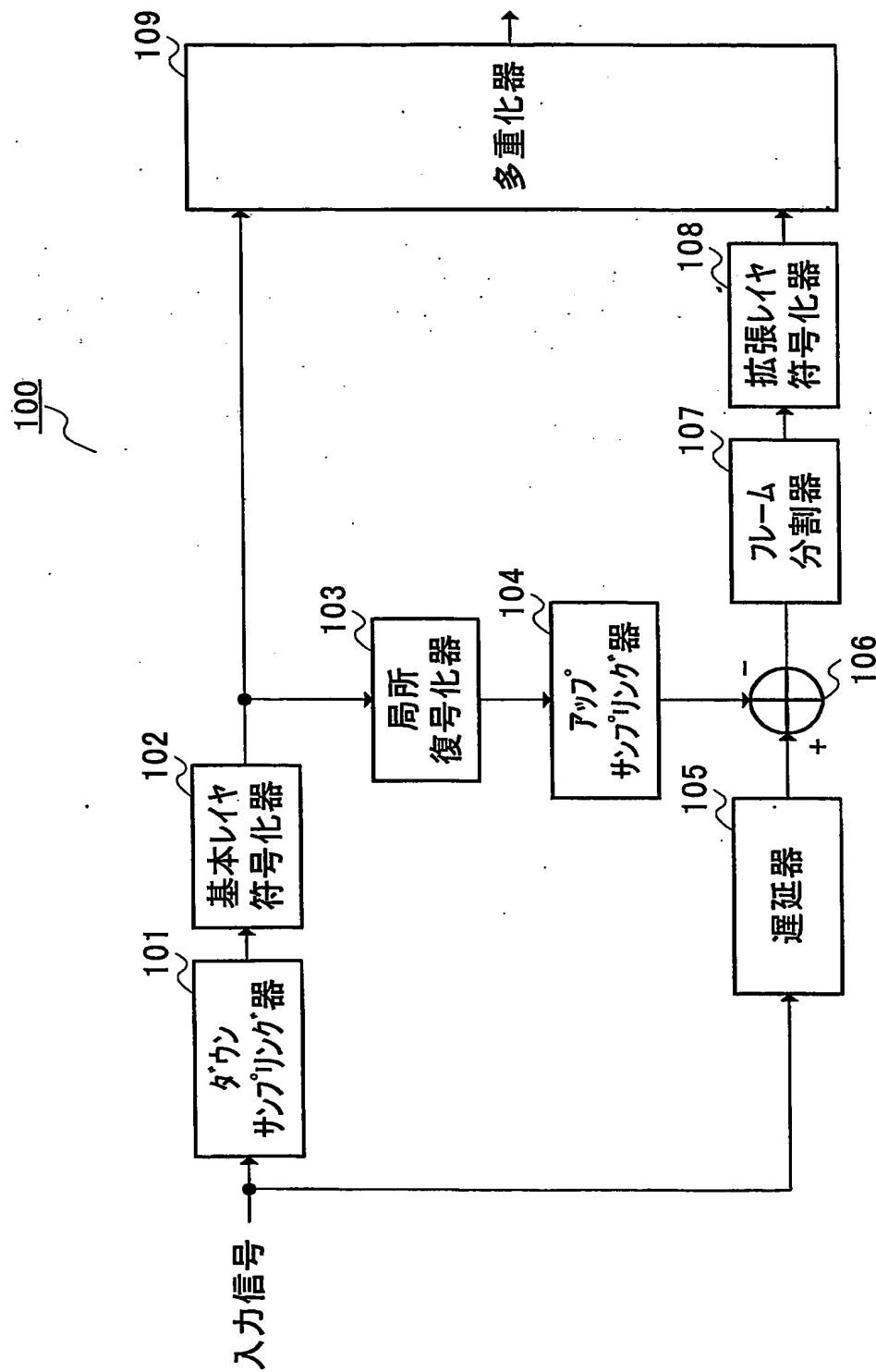


図3



4/26

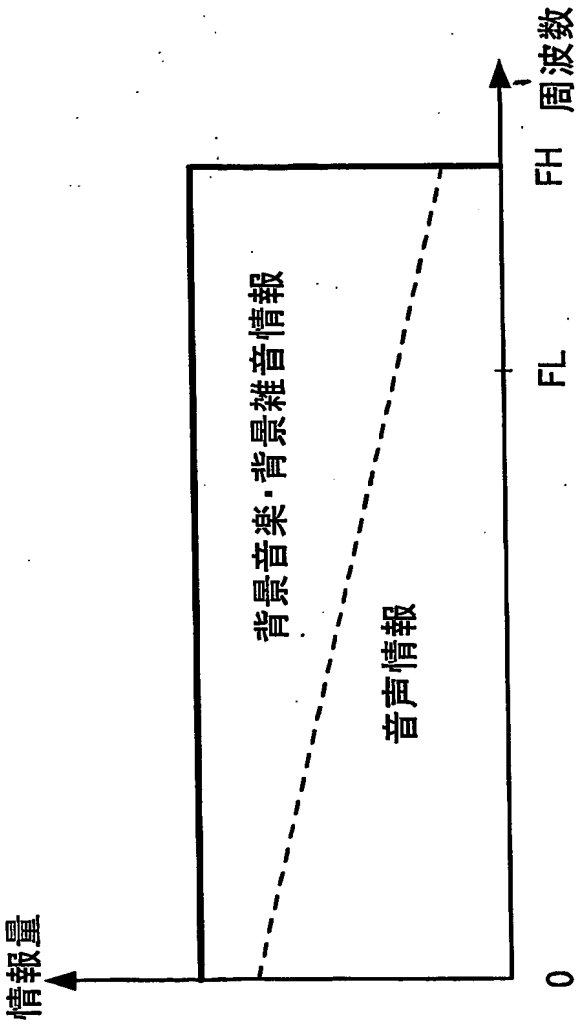


図4

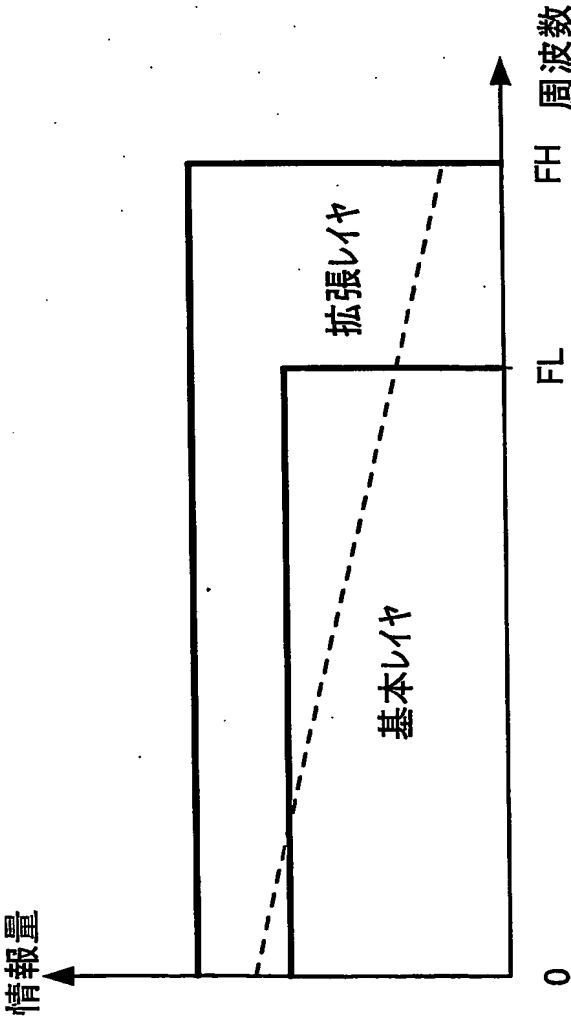


図5

6/26

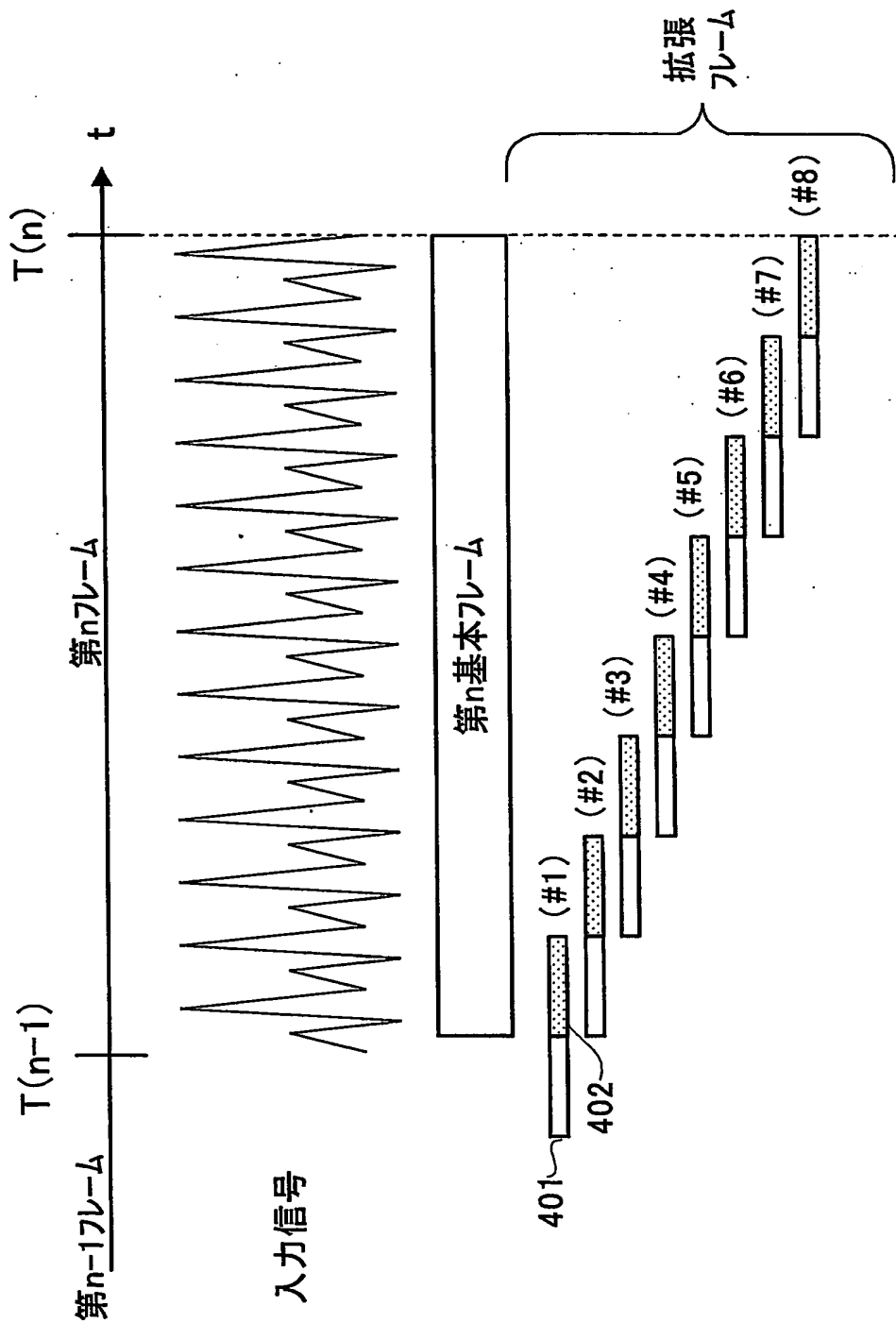


図6

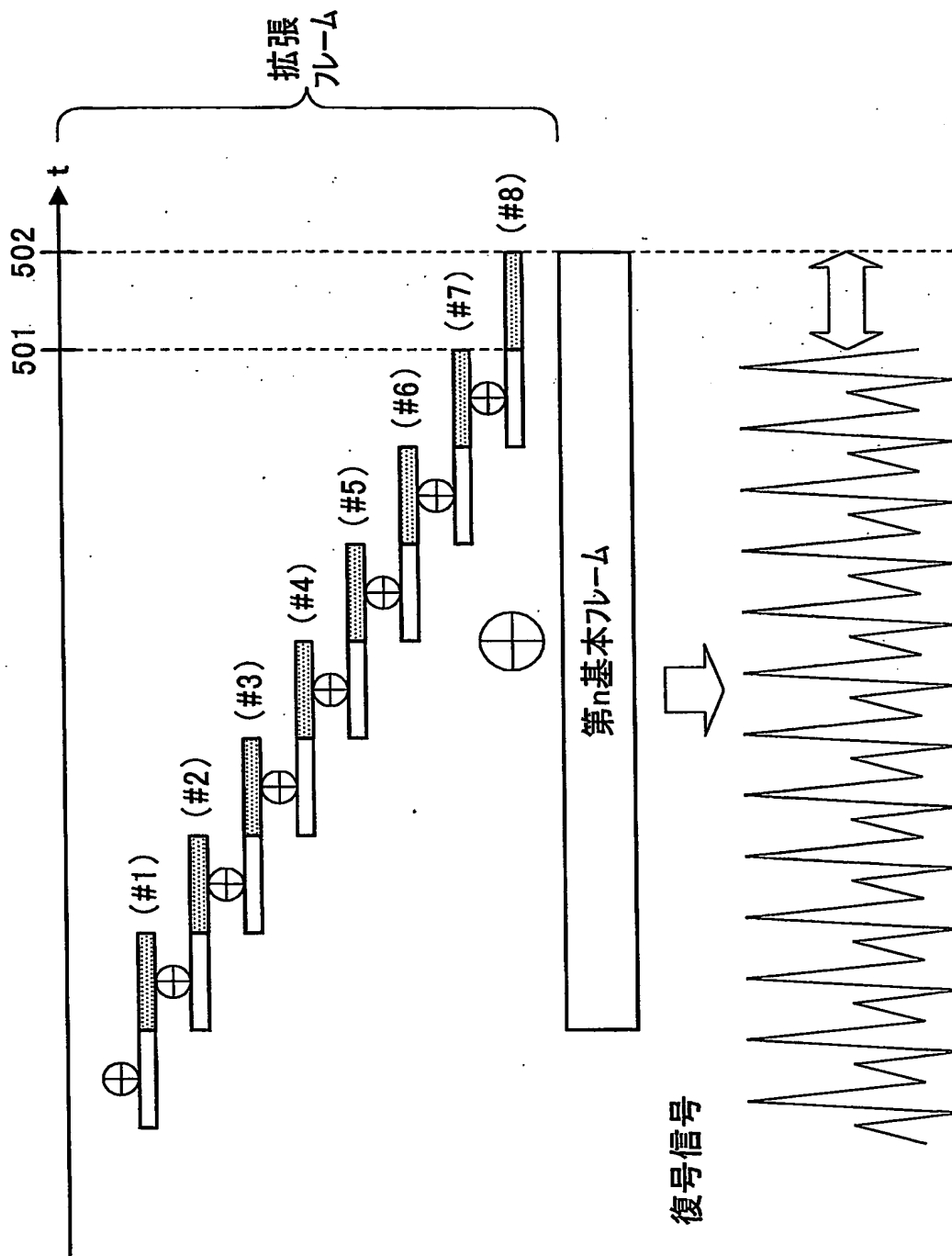


図7

8/26

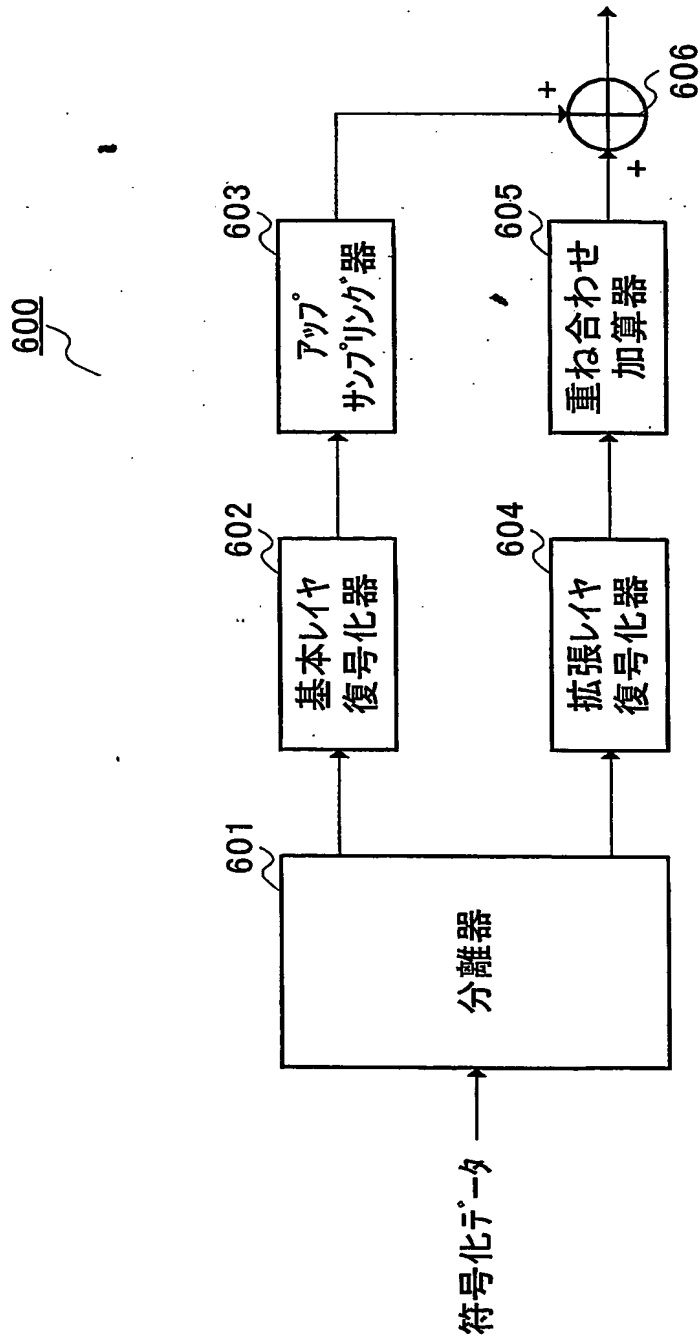


図8

9/26

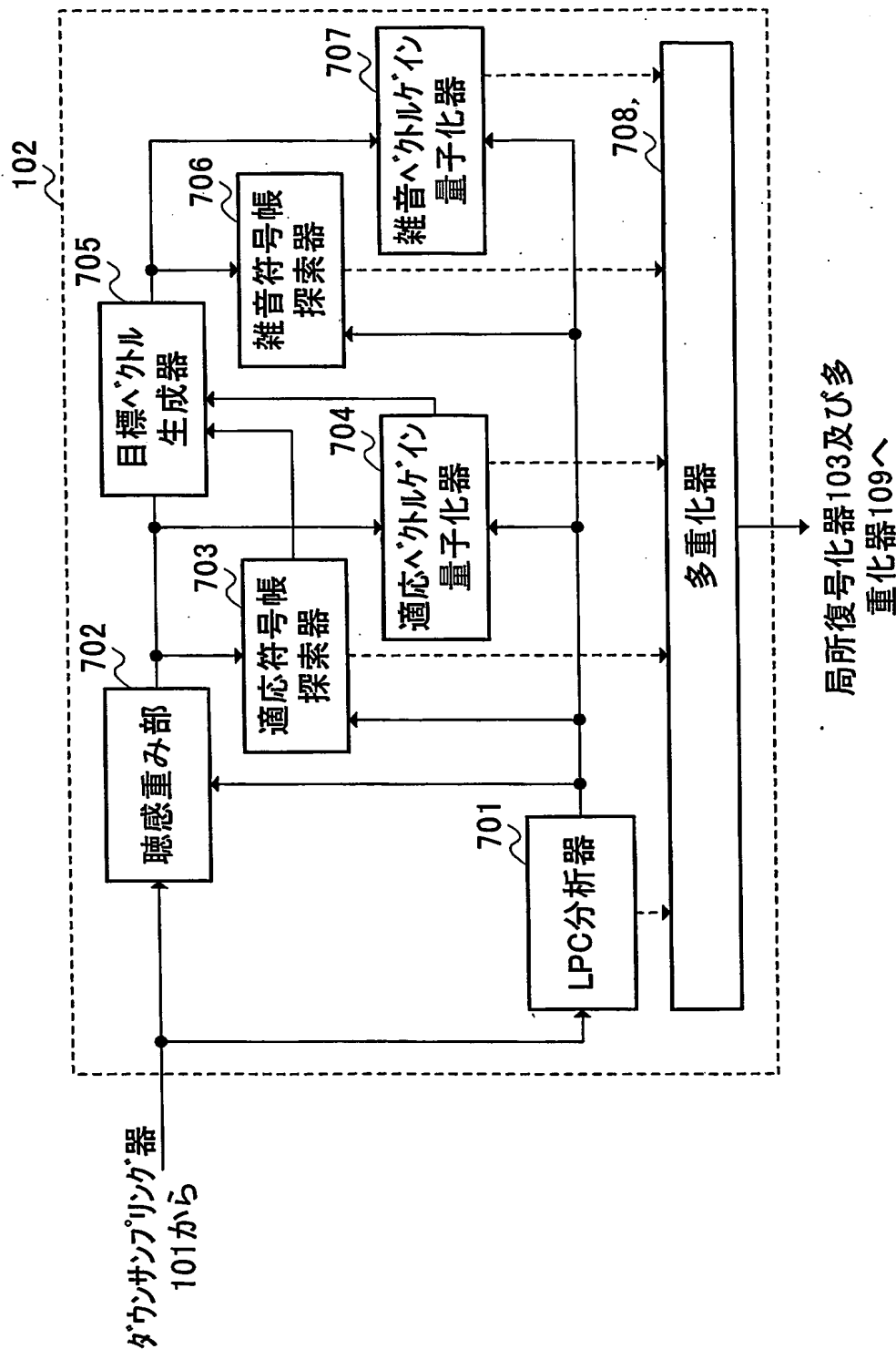


図9

10/26

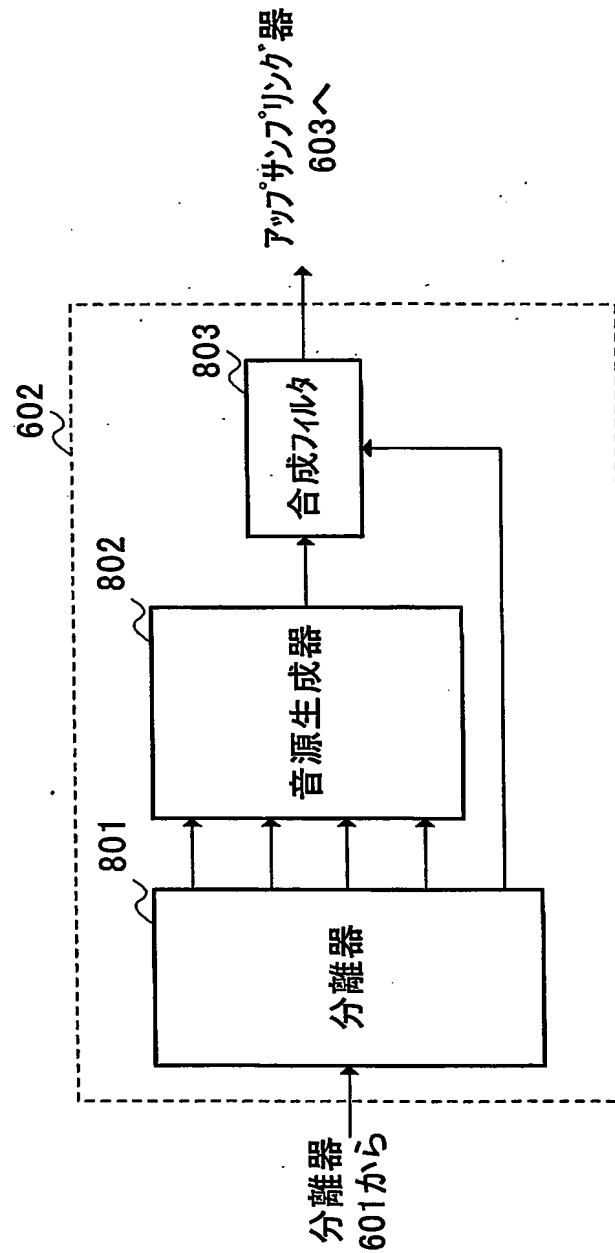


図10

11/26

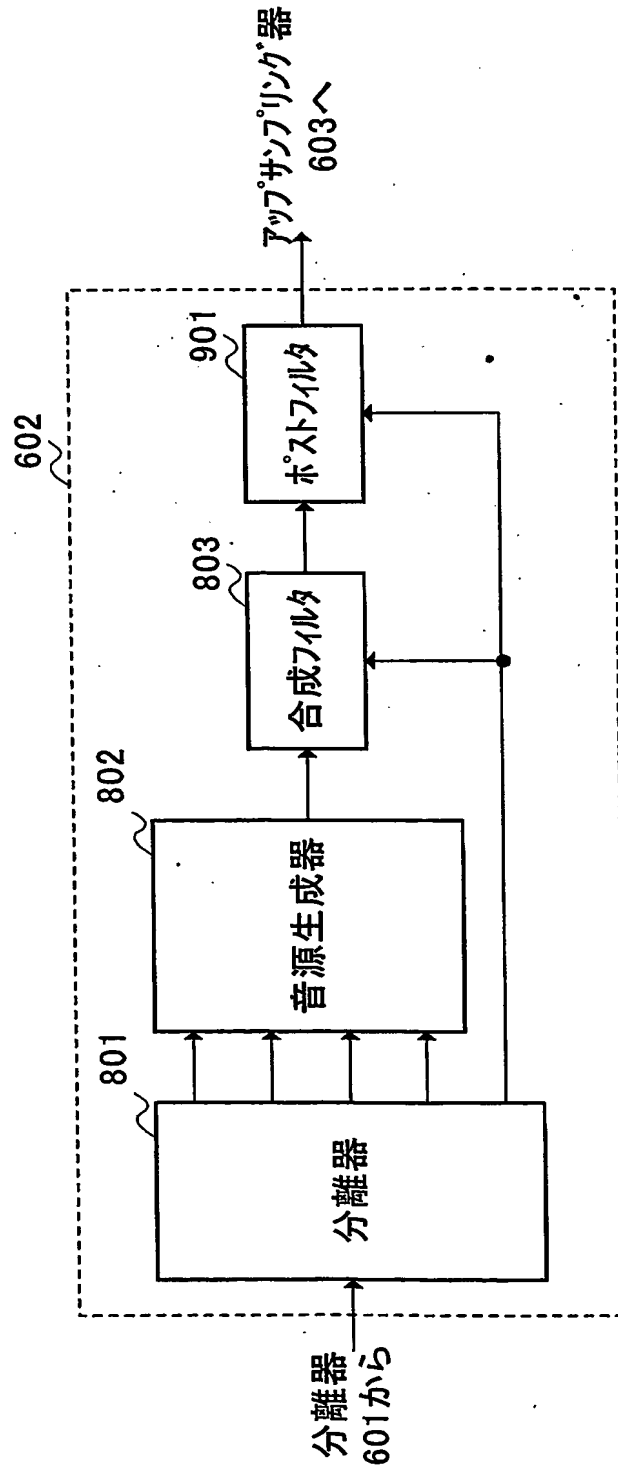


図11



12/26

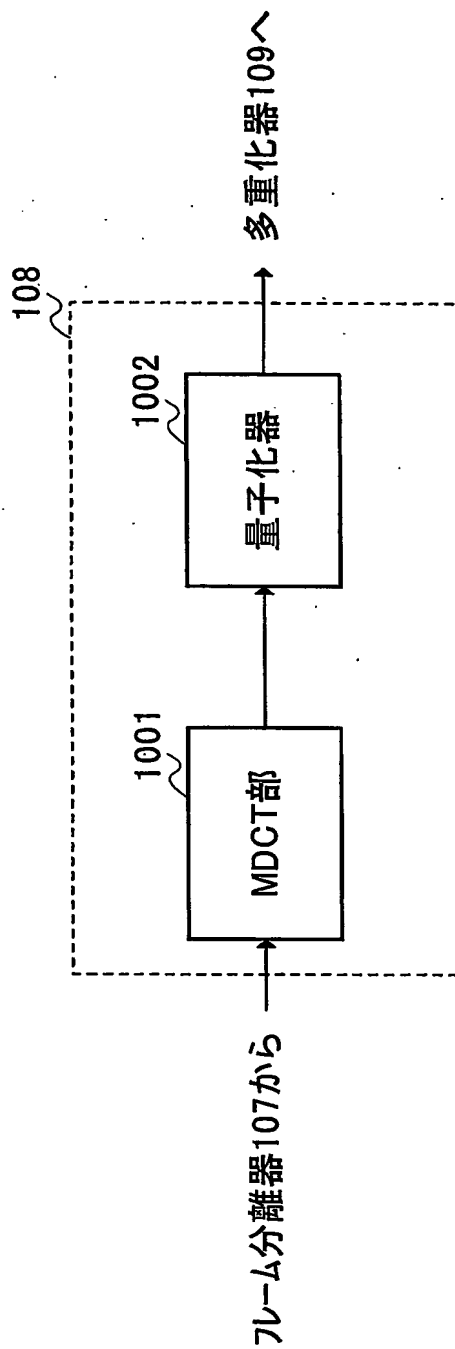


図12

13/26

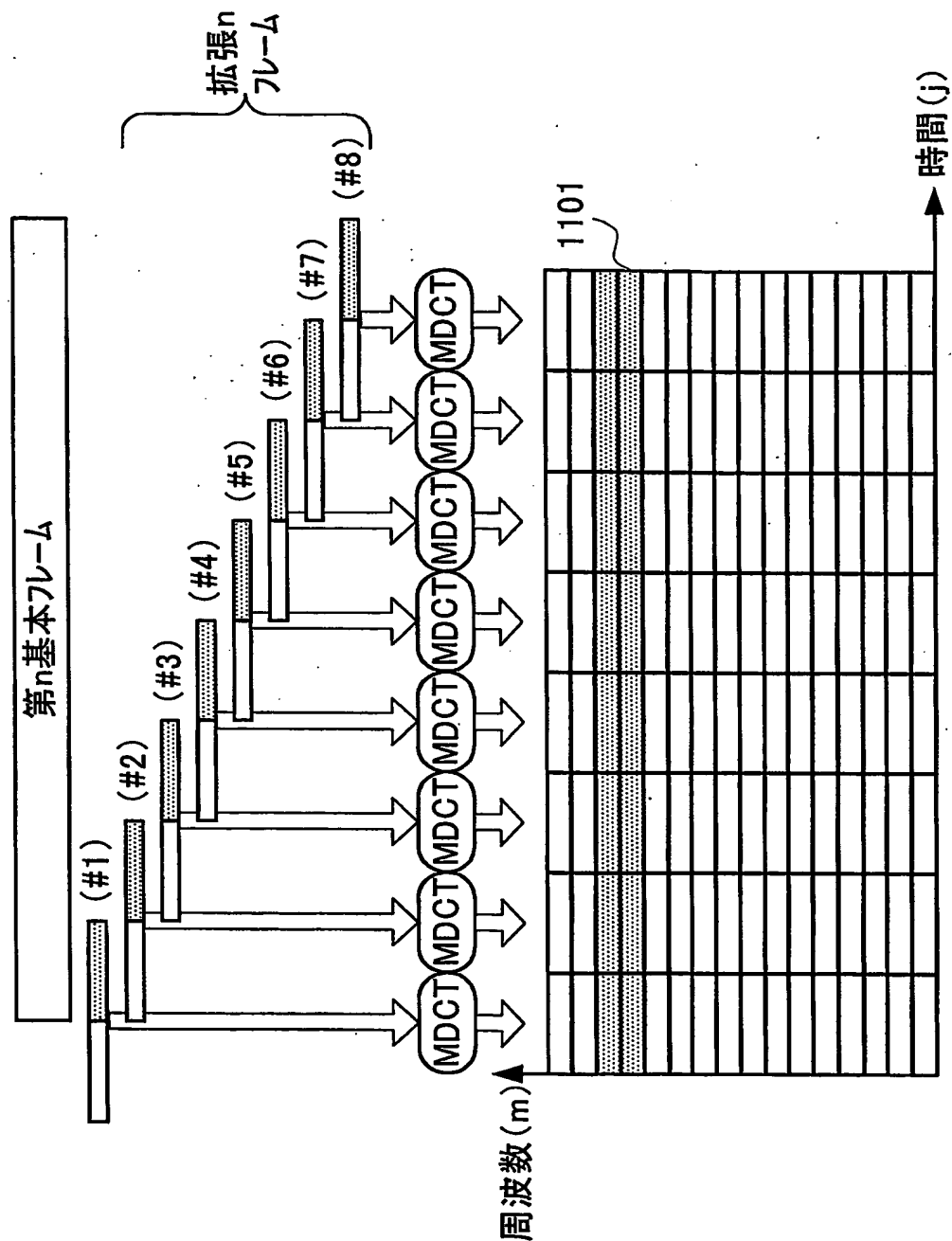


図13

14/26

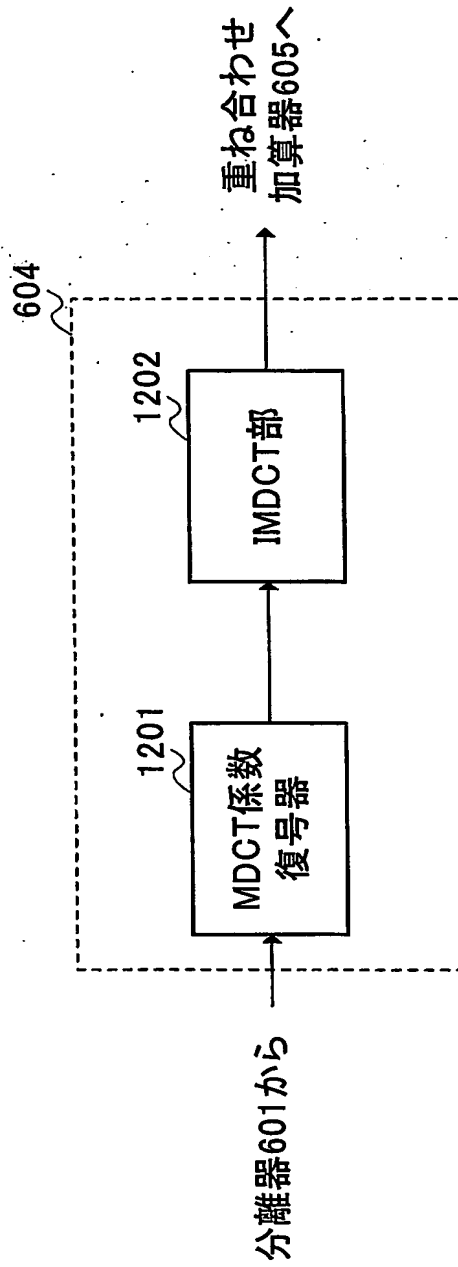


図14

15/26

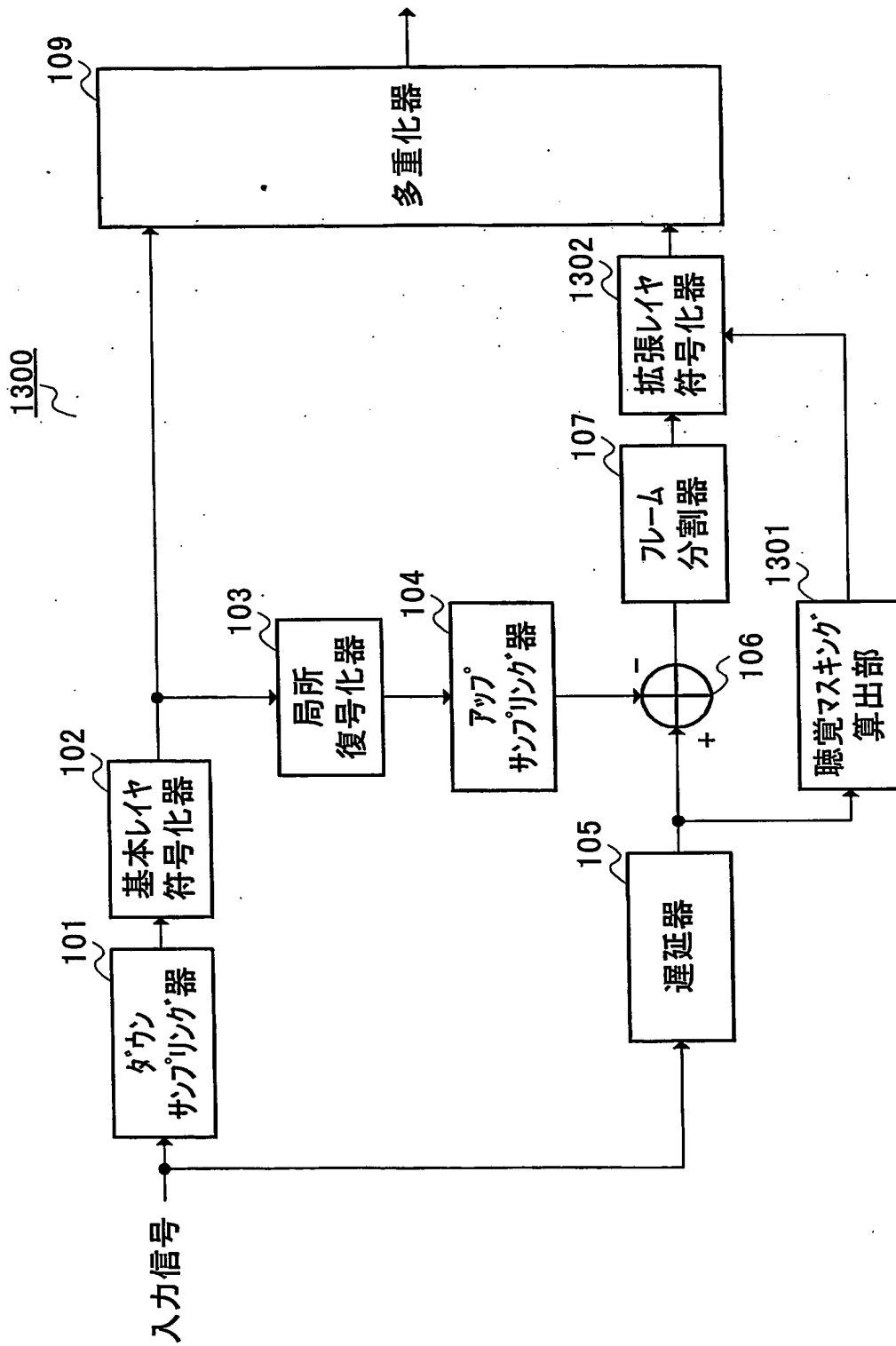
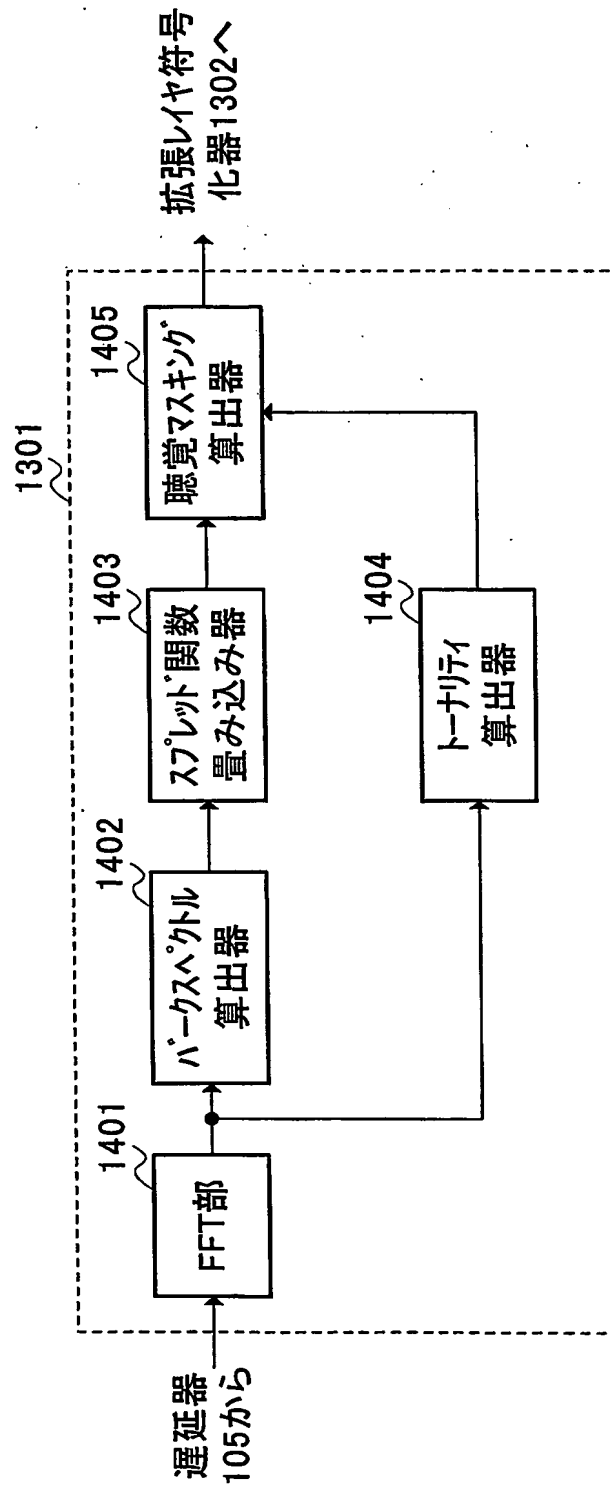


図15



16 図

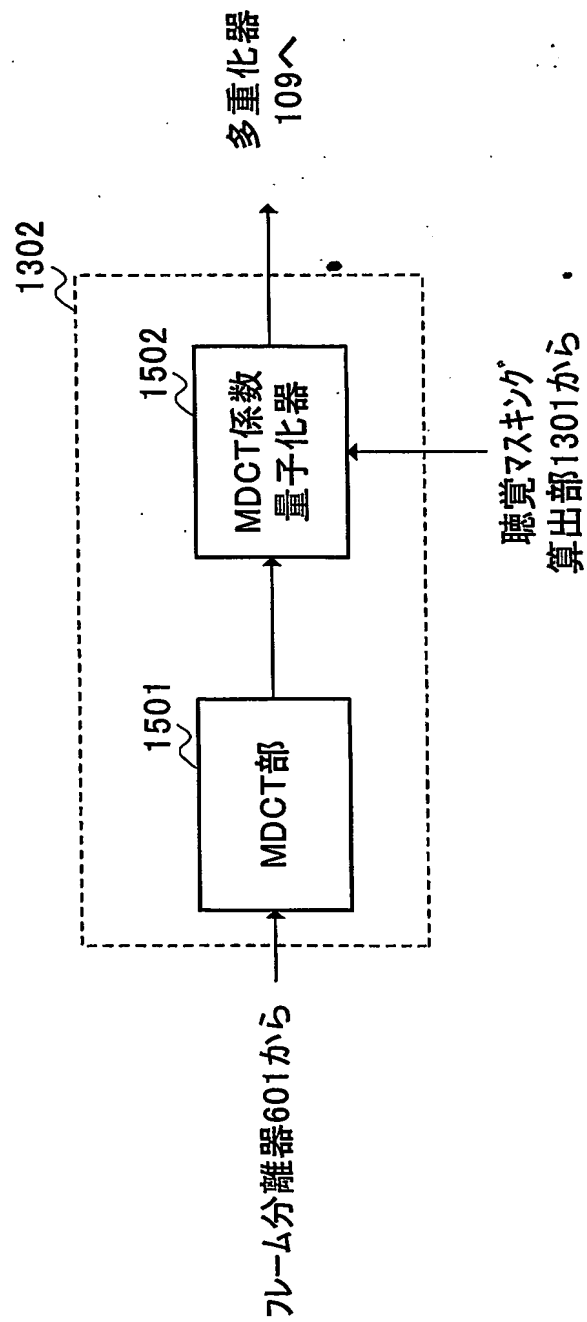


図17

18/26

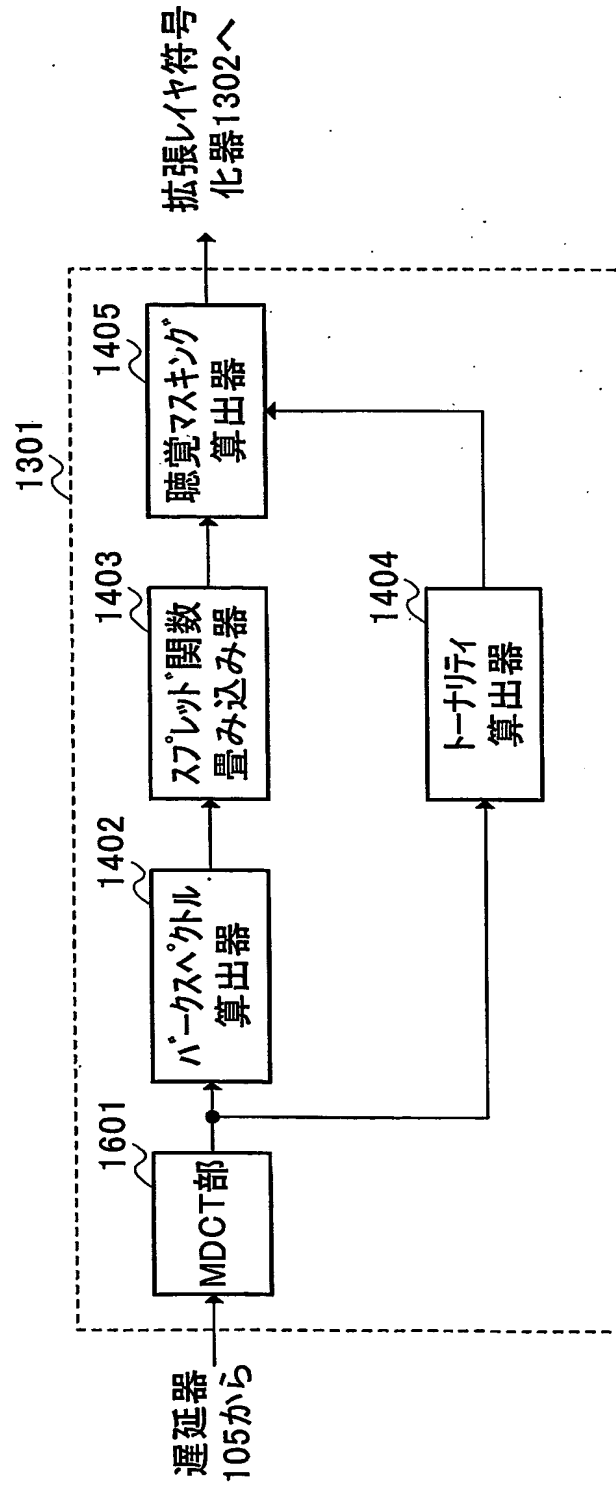


図18

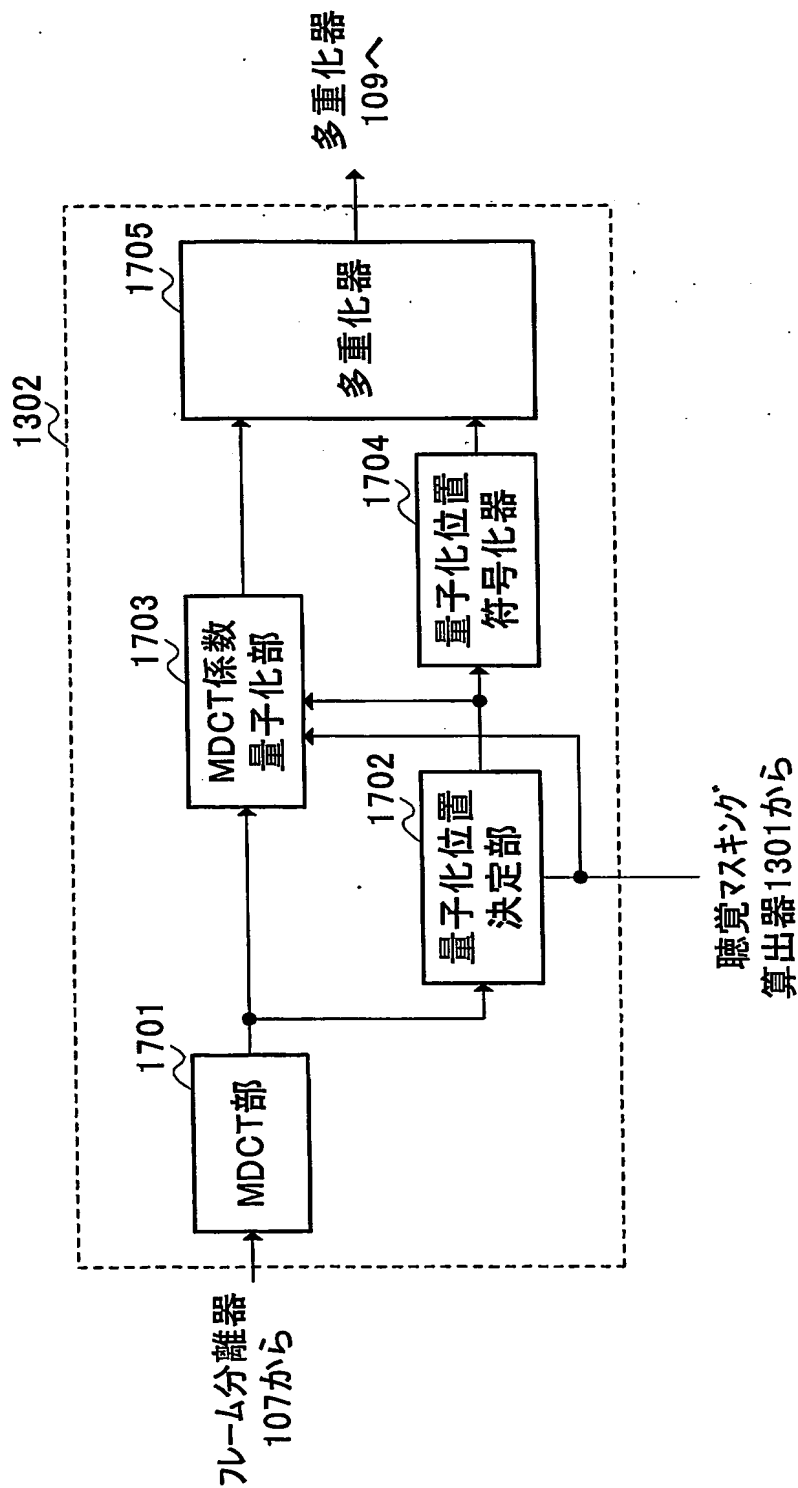


図19



20/26

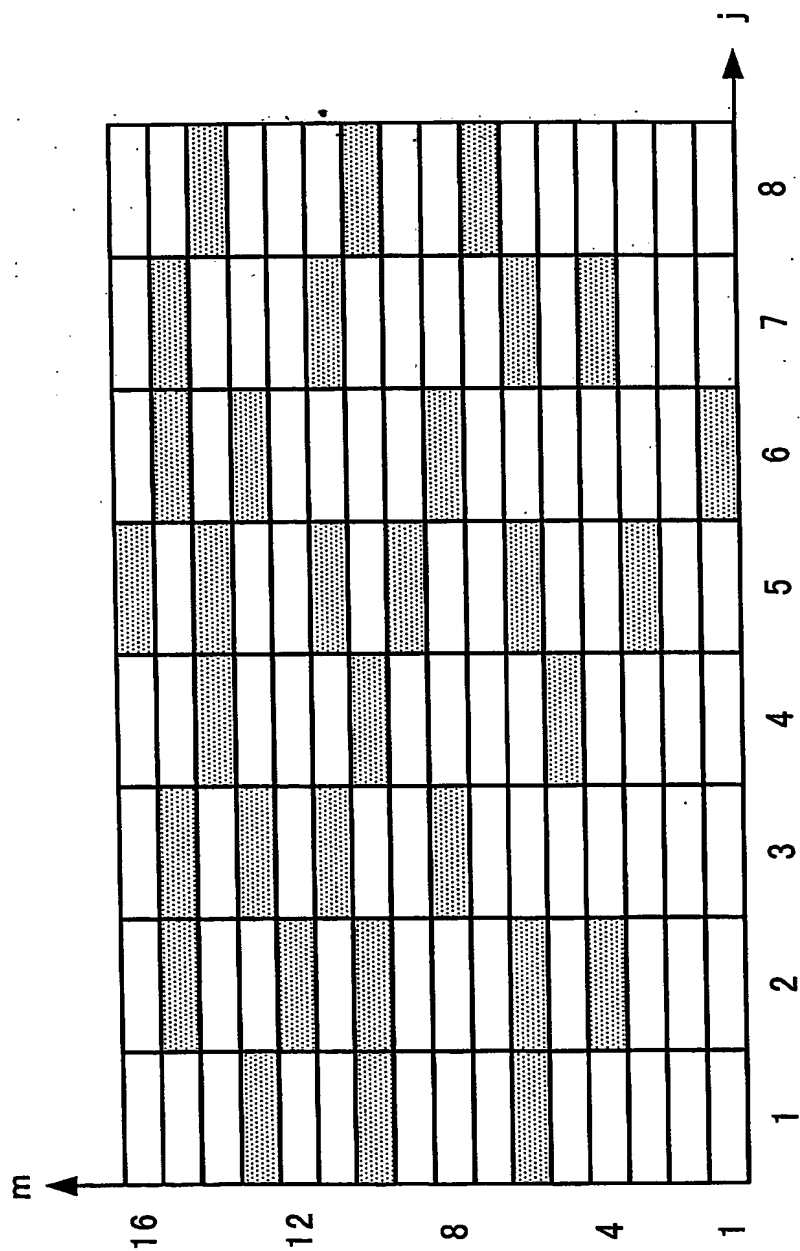


図20

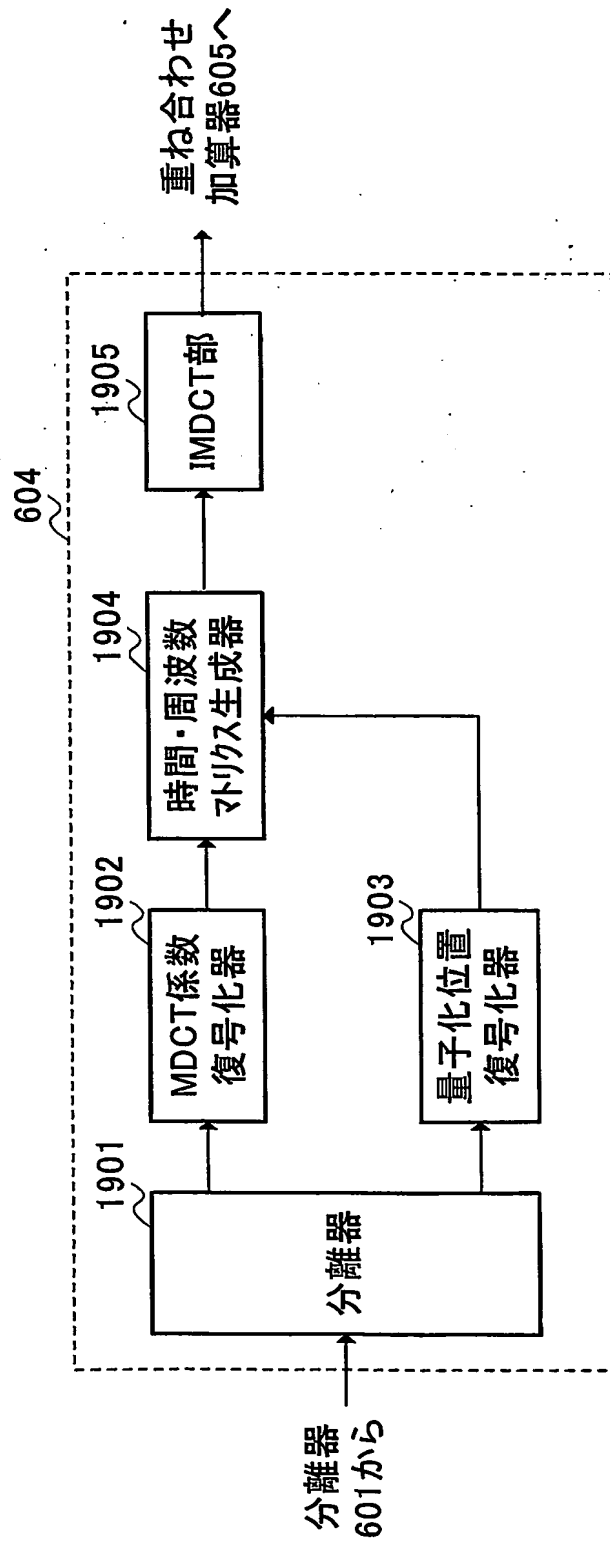


図21

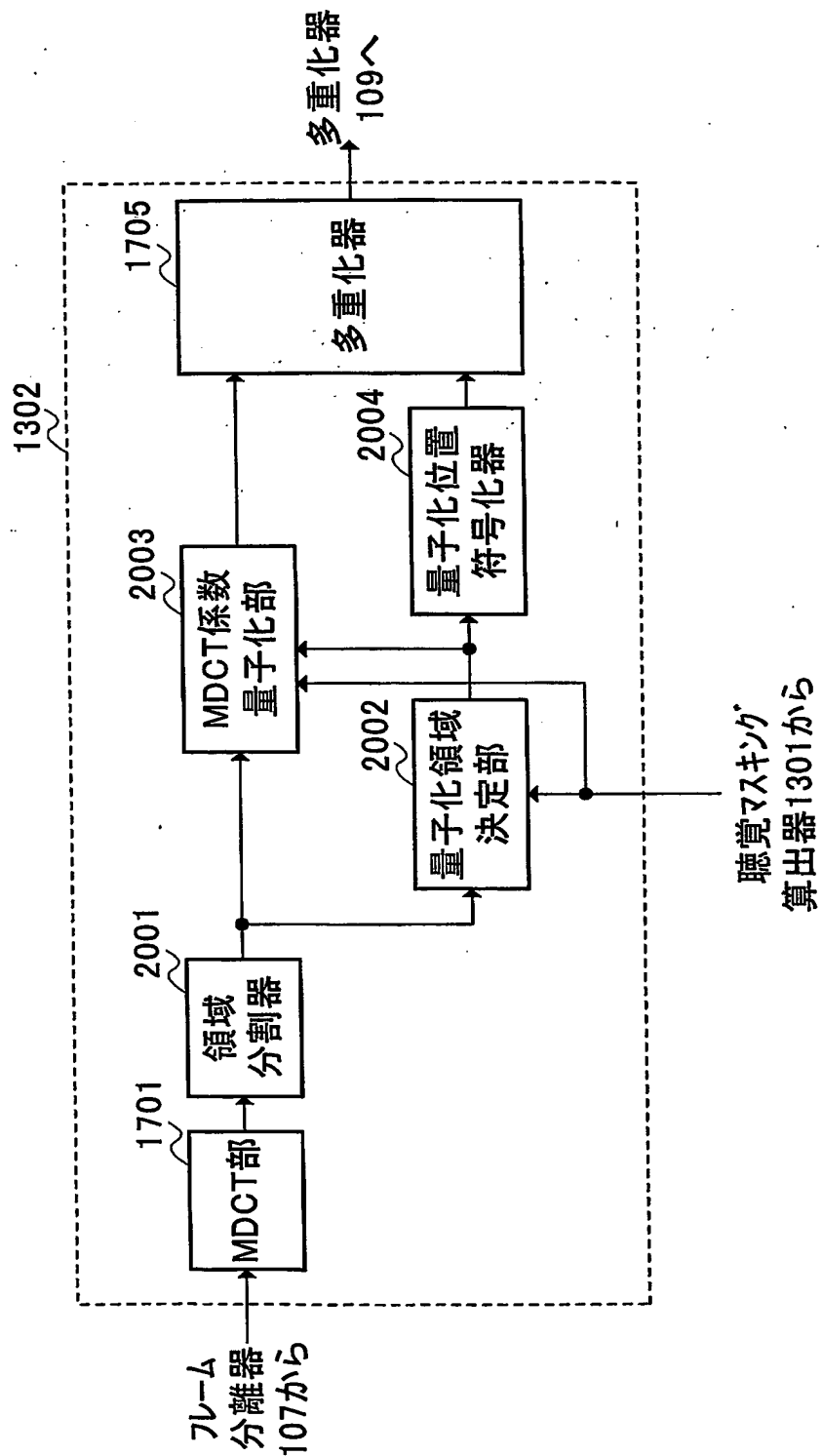


図22

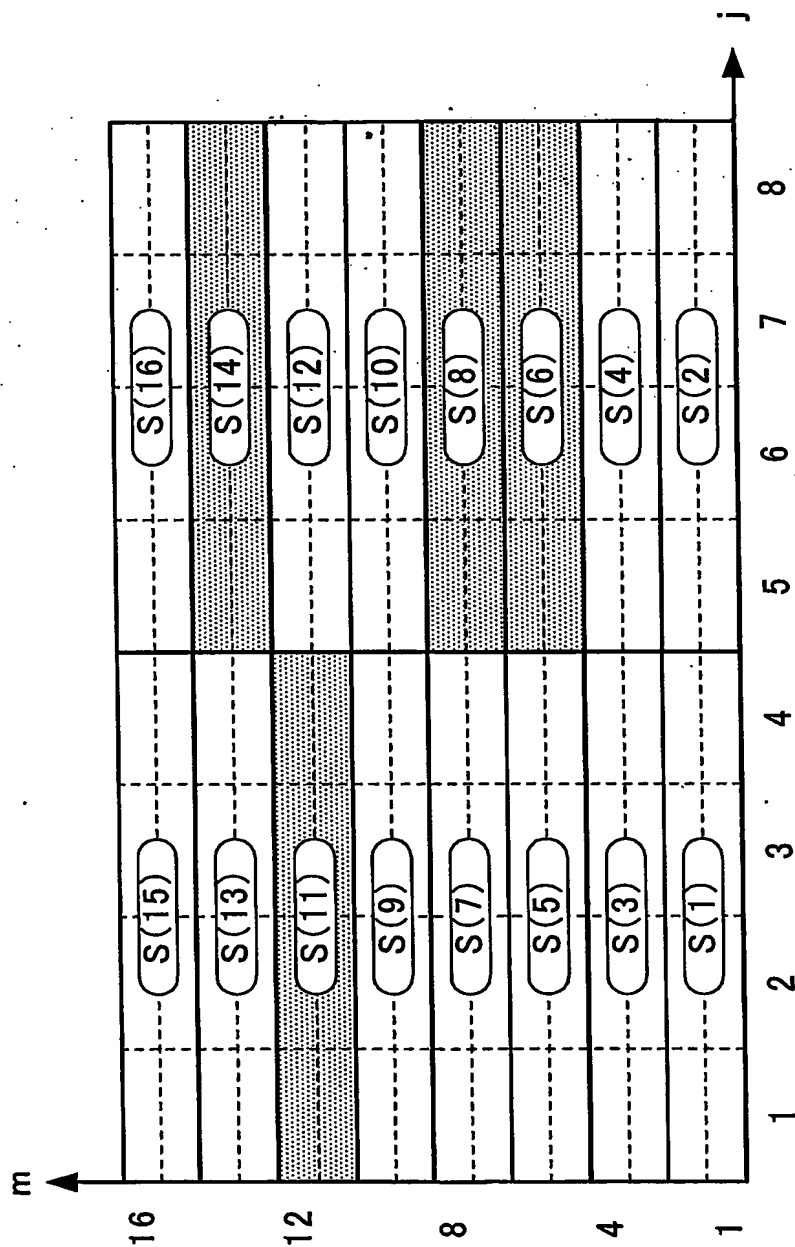


図23

24/26

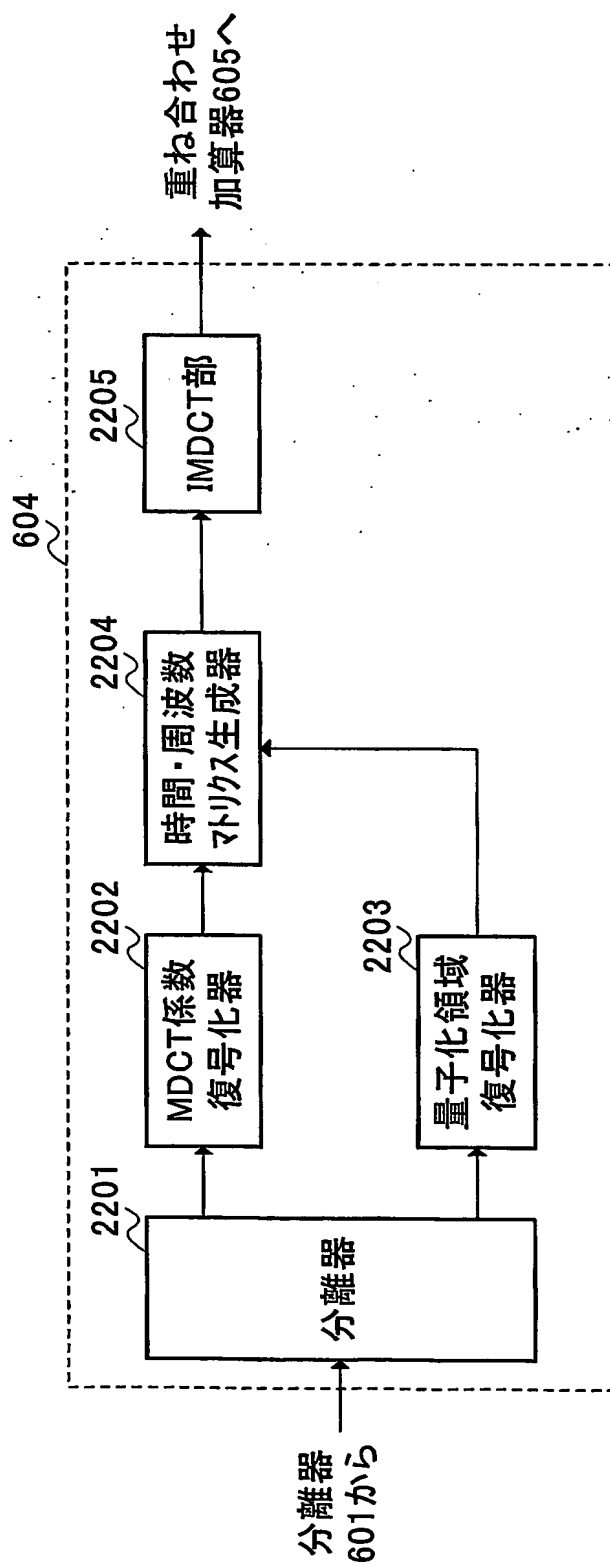


図24

25/26

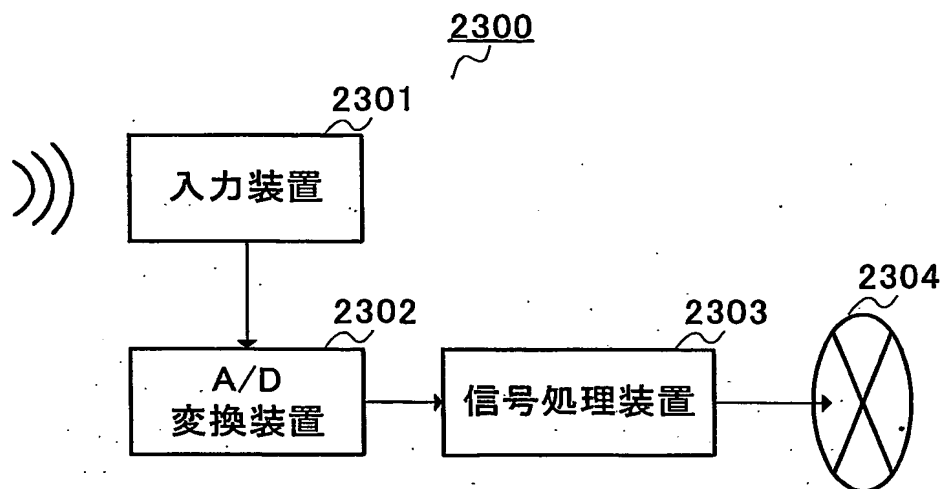


図25

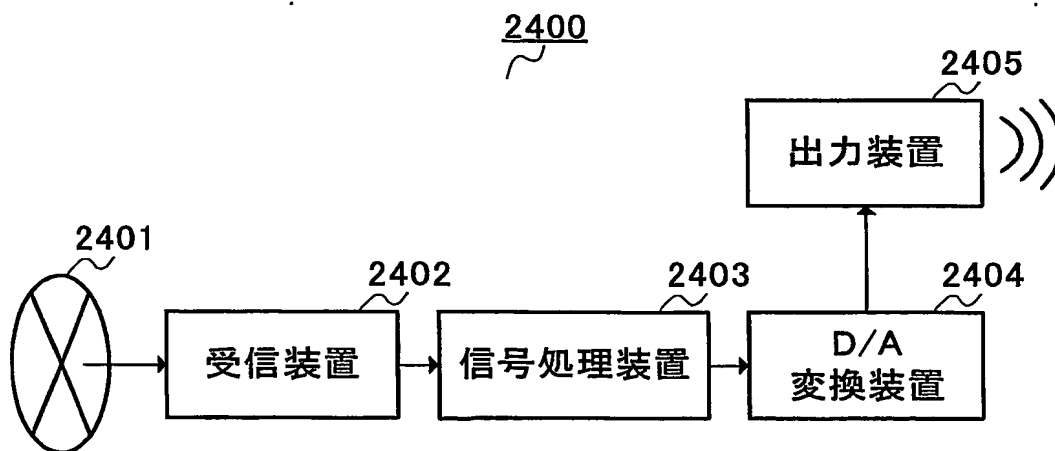


図26

26/26

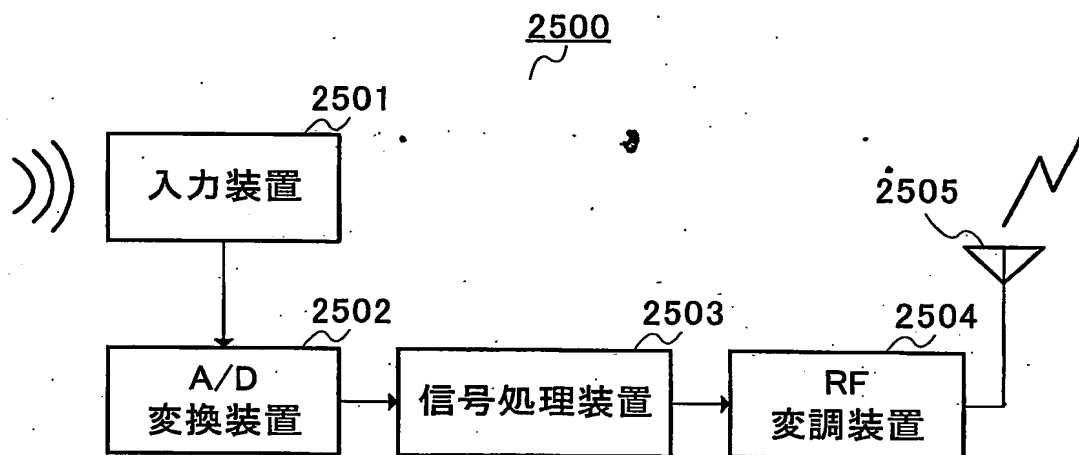


図27

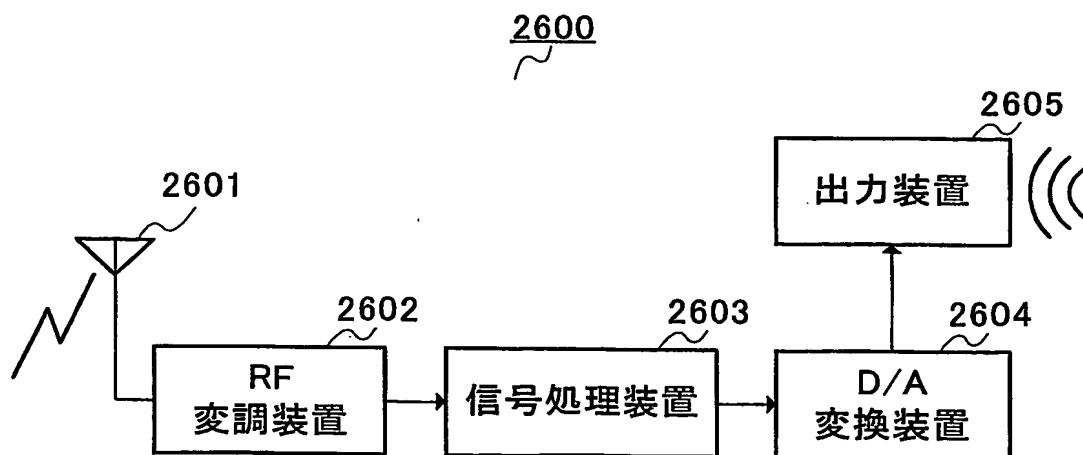


図28

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP03/10247

## A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl<sup>7</sup> G10L19/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl<sup>7</sup> G10L19/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Toroku Jitsuyo Shinan Koho	1994-2003
Kokai Jitsuyo Shinan Koho	1971-2003	Jitsuyo Shinan Toroku Koho	1996-2003

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	JP 8-263096 A (Nippon Telegraph And Telephone Corp.), 11 October, 1996 (11.10.96), (Family: none)	1-7, 10-13, 16-20 8, 9, 14, 15
Y A	JP 2000-3193 A (NEC Corp.), 07 January, 2000 (07.01.00), Column 25, Par. Nos. [0110] to [0120] & WO 99/66497 A1 & EP 1087378 A1	1-7, 10-13, 16-20 8, 9, 14, 15
Y A	JP 9-127996 A (Sony Corp.), 16 May, 1997 (16.05.97), & US 575222 A1 & EP 770988 A1	1-7, 10-13, 16-20 8, 9, 14, 15
Y A	JP 10-285046 A (Sony Corp.), 23 October, 1998 (23.10.98), (Family: none)	4-7, 13 8, 9, 14, 15

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

\* Special categories of cited documents:  
 "A" document defining the general state of the art which is not considered to be of particular relevance  
 "E" earlier document but published on or after the international filing date  
 "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)  
 "O" document referring to an oral disclosure, use, exhibition or other means  
 "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  
 "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone  
 "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art  
 "&" document member of the same patent family

Date of the actual completion of the international search  
02 September, 2003 (02.09.03)

Date of mailing of the international search report  
16 September, 2003 (16.09.03)

Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.



## A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl<sup>7</sup> G10L19/00

## B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl<sup>7</sup> G10L19/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年  
 日本国公開実用新案公報 1971-2003年  
 日本国登録実用新案公報 1994-2003年  
 日本国実用新案登録公報 1996-2003年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

## C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	J P 8-263096 A (日本電信電話株式会社) 1996.10.11 (ファミリーなし)	1-7, 10-13, 16-20
A		8, 9, 14, 15
Y	J P 2000-3193 A (日本電気株式会社) 2000.01.07, 第25欄【0110】-【0120】, & W099/66497 A1 & EP 1087378 A1	1-7, 10-13, 16-20
A		8, 9, 14, 15
Y	J P 9-127996 A (ソニー株式会社) 1997.05.16 & US 575222 A1 & EP 770988 A1	1-7, 10-13, 16-20
A		8, 9, 14, 15

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

## \* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的な技術水準を示すもの  
 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの  
 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)  
 「O」 口頭による開示、使用、展示等に関する文献  
 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&amp;」 同一パテントファミリー文献

国際調査を完了した日

02.09.03

国際調査報告の発送日

16.09.03

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

渡邊 聡

5C

8622

電話番号 03-3581-1101 内線 3540

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	JP 10-285046 A (ソニー株式会社) 1998.10.23 (ファミリーなし)	4-7, 13
A		8, 9, 14, 15

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**